Doctoral theses at NTNU, 2019:212

Espen Dyrnes Stabell

Deep Sea Uncertainty

Studies in Environmental Ethics and Decision-Making

NTNU
Norwegian University of
Science and Technology

Doctoral thesis

Espen Dyrnes Stabell

ISBN 978-82-326-4020-1 (printed ver.)
ISBN 978-82-326-4021-8 (electronic ver.)
ISSN 1503-8181

NTNU
Norwegian University of
Science and Technology

Thesis for the Degree of
Philosophiae Doctor
Faculty of Humanities
Department of Philosophy and Religious Studies

NTNU
Norwegian University of Science and Technology

Doctoral theses at NTNU, 2019:212

NTNU

Espen Dyrnes Stabell

Deep Sea Uncertainty

Studies in Environmental Ethics and Decision-Making

Thesis for the Degree of Philosophiae Doctor

Trondheim, August 2019

Norwegian University of Science and Technology
Faculty of Humanities
Department of Philosophy and Religious Studies

NTNU
Norwegian University of
Science and Technology

NTNU
Norwegian University of Science and Technology

Thesis for the Degree of Philosophiae Doctor

Faculty of Humanities
Department of Philosophy and Religious Studies

ISBN 978-82-326-4020-1 (printed ver.)
ISBN 978-82-326-4021-8 (electronic ver.)
ISSN 1503-8181

Doctoral theses at NTNU, 2019:212

Printed by NTNU Grafisk senter

Contents

Preface

The Pilot Programme on Deep Sea Mining

The research for this dissertation was conducted as part of an interdisciplinary research project at the Norwegian University of Science and Technology (NTNU), called the Pilot Programme on Deep Sea Mining (referred to hereafter as the Pilot).[1] The Pilot is a collaboration between the Faculty of Humanities, the Faculty of Information Technology and Electrical Engineering, the Faculty of Engineering, and the Faculty of Natural Sciences. It is part of the NTNU strategic area NTNU Oceans and includes research in geology, geophysics, engineering, history, environmental studies and philosophy. An aim of the Pilot is to develop 'new solutions for evaluation, exploration and extraction of sea-based minerals'. Moreover, the Pilot seeks to investigate impacts on the environment and society. The 'ethics and social responsibility' part of the project, which the current dissertation falls under, is meant to look at 'normative questions involved in the development of new technology in general, and within deep sea mining in particular'.

Deep sea mining, which is the process of retrieving mineral deposits from the deep sea, is not yet an established commercial practise. Why is it important to discuss it from an ethical point of view at this early stage? As pointed out by David Collingridge (1980), policy decisions about new technologies involve a dilemma. At an early stage of research and development of a new technology, we cannot know for certain what its most important impacts will be. Hence, we cannot know which ethical issues are most important to address, or even in some cases what the issues are. However, if we wait and see what the impacts are, it is often too late to do anything about them. For instance, at the dawn of the industrial revolution, no one foresaw the effects that the rise in greenhouse gas emissions from major industries would have on the global climate.

[1] Webpage: https://www.ntnu.edu/oceans/deep-sea-mining (accessed 01.04.2019). Quotations in this paragraph are from the webpage.

When scientists in the 20th century discovered the connection between anthropogenic emissions and climate change, however, some climate change was already underway and could not be stopped.

The Pilot was formed on the background of the influential Responsible Research and Innovation (RRI) paradigm for interdisciplinary research.[2] A basic idea behind RRI is that when we inevitably encounter the quandary of Collingridge's dilemma, it is best to err on the side of precaution: issues of ethical, social or political importance should be investigated and, if possible, addressed at an early stage in researching and developing new technologies, even if there is uncertainty about their impacts. This provides an important rationale for including research in philosophy and other humanistic and social science disciplines in predominantly technological projects such as the Pilot (Lucivero, Swierstra, and Boenink 2011).

Acknowledgements

First of all, I thank my supervisors at NTNU, Siri Granum Carson and Bjørn Kåre Myskja, for their guidance and support, and for incisive and constructive feedback on all portions of the manuscript. In 2017, I spent two months at the University of British Columbia (UBC), W. Maurice Young Centre for Applied Ethics, working with Daniel Steel on what was to become the third study of the dissertation. After my stay, I was fortunate to get Dan on board as co-supervisor for my thesis. A special thanks to Dan for believing in the project and for taking the time to co-write the study with me, and to Allen Alvarez, Itai Bavli, Susan Cox and everyone at the Centre and the philosophy departments at UBC and Simon Fraser University for making my stay in Vancouver such a productive and memorable experience.

[2] The influence of RRI as a research paradigm is indicated by its centrality to the huge EU research and innovation programme Horizon 2020 (https://ec.europa.eu/programmes/horizon2020/en/h2020-section/responsible-research-innovation (accessed 01.10.2018)), and to national institutions such as the Research Council of Norway (https://www.forskningsradet.no/prognett-biotek2021/Ansvarlig_forskning_og_innovasjon_RRI/1254026368408 (accessed 01.10.2018)).

I would also like to thank Steinar Ellefmo and the participants in the Pilot, as well as the director of NTNU Oceans, Ingrid Schjølberg, for giving me the opportunity to work with researchers in history, geology, engineering and several other fields in an exciting and innovative interdisciplinary project. The dissertation has benefitted greatly from discussions with bright minds from the Pilot and affiliated projects.

The papers collected in this dissertation have been presented at various seminars and conferences in several countries. I am grateful to everyone who provided feedback, criticism and insightful commentary, including members of VERP (currently administered by Hermann Køhn Sæther and Mons Andreas Nyquist) at the Department of Philosophy and Religious Studies at NTNU, the Risk and Safety Seminar at the Department of Philosophy at KTH Royal Institute of Technology, the Practical Philosophy Group at the University of Oslo, as well as Eric Brandstedt, Carlos Joly, Sven R. Nyholm, Robert Sparrow, Helga Varden and Jonathan Wolff. Especially, I thank Morten Langfeldt Dahlback for excellent comments on earlier drafts of the manuscript, Sigurd Hverven for close readings of drafts and rewarding discussions about environmental philosophy, and Christian Munthe for extensive and helpful commentary on the introductory chapter and the third study of the dissertation.

Finally, I thank Camilla, Lykke and Hedda for patiently putting up with an absent-minded and moody husband and father, and the rest of my family for love and support. A special thanks to my mother, Anne-Mette Dyrnes, who also read and commented on the manuscript. The dissertation is dedicated to the memory of my father, Harald Stabell.

Introduction

Preliminaries

This dissertation discusses philosophical issues related to environmental decision-making, with a particular focus on the case of *deep sea mining*. Deep sea mining involves retrieving mineral deposits from the ocean floor at great depths in ecologically complex environments. This form of mining is still in its infancy, with high expectations for societal benefits and economic gain. At the same time, there are significant uncertainties regarding environmental impact and how to mitigate potential harm. The dissertation focuses on three main research questions regarding the ethics of deep sea mining:

1. Is 'existence value' – understood as the value humans ascribe to the existence of something, regardless of whether it is or will be of any particular use to them – important in evaluating activities that come with a risk of species extinction?
2. How should choices be made in environmental cases where, given a set of options, such as alternative courses of action, the options are not considered better or worse than each other with respect to the relevant criteria, nor are they considered equally good?
3. If precautionary measures are to be implemented against environmental threats, how can we ensure a fair distribution of the costs of such measures?

The first question is relevant to deep sea mining because, as will be explained in more detail below, deep sea mining involves a risk of species extinction. The second question is relevant because in evaluating our options regarding deep sea mining, we may face choices between options that are difficult to compare and evaluate on the grounds of their respective merits and

demerits. The third is relevant because precautionary measures are suggested by many authors writing on deep sea mining, and are likely to be required by regulators in deep sea mining projects.

I conclude the dissertation with a discussion of whether deep sea mining should be conducted or not. In the current introductory chapter, I explain the empirical, theoretical and methodological background of the dissertation. I begin by providing background information about deep sea mining, focusing especially on environmental and societal issues relevant to my studies. I then explain my use of the concepts of 'risk' and 'uncertainty' as well as the role of the 'precautionary principle' in the dissertation, and situate my studies in the environmental ethics debate. Finally, I lay out the more general philosophical (theoretical and methodological) framework of the dissertation, and summarize the main arguments and conclusions of the studies. Topics for further research are suggested at the end of the dissertation.

Deep sea mining

Deep sea mining is a mineral retrieving process that takes place in the deep sea – the area of the ocean below 200 meters, which covers about 65% of the Earth's surface. Deep sea mining has become particularly important in recent years as a potential source of raw materials. After many decades of research on and interest in seabed mineral resources, technological developments have allowed potential exploitation of previously unreachable deposits. Increasing demand, combined with a scarcity of some minerals, has triggered renewed attention for the possibility of exploiting deep sea resources (ECORYS 2014).

To date, more than 2.5 million square kilometres of the seafloor are under contract for mineral exploration, and the number is rapidly increasing.[3] The development is driven by global

[3] http://www.mdpi.com/journal/minerals/special_issues/marine_minerals_gas (accessed 01.10.2018). At least 28 contracts have been given by the International Seabed Authority (ISA) for exploration in international waters (the 'Area'), see https://www.isa.org.jm/deep-seabed-minerals-contractors (accessed 01.03.2018). Moreover, several projects have been proposed and are getting ready to start up in exclusive economic zones (within national

metal markets, technological innovation, and the establishment of regulatory frameworks by coastal nations and the International Seabed Authority (ISA). Although much of the technology has not yet been proven on great depths, technologies now exist for the mining of seafloor massive sulphides, manganese nodules, and phosphorite, and is in the final stages of development for cobalt-rich ferromanganese crusts. It has been suggested that within the next few years, the first deep-ocean mines will have begun operations and a new industry will have seen the light of day.[4]

Concerns have been raised about the impacts of deep sea mining on the environment (e.g. Halfar and Fujita 2002, ECORYS 2014, Van Dover et al. 2017). At present, our knowledge of how deep-sea benthic ecosystems and the functions they serve respond to human pressures is very limited (ECORYS 2014).[5] Nevertheless, the current state of knowledge gives a general picture of plausible environmental impacts.

As with all mining activity, deep sea mining will directly impact habitats, resulting in the removal of fauna and seabed rock and minerals. According to some authors, deep sea mining will impose disturbances 'with novel nature, strength, stability and spatial and temporal scale on deep-sea ecosystems' (Gollner et al. 2017, see also Van Dover 2014, Ramirez-Llodra et al. 2011, Hauton et al. 2017, Boschen et al. 2013, Van Dover 2010).

Exploration licenses are being granted for three different types of deep-sea mineral resources (Gollner et al. 2017):

1. polymetallic (seafloor massive) sulphides on active hydrothermal vents – which are fissures in the earth's surface from which geothermally heated water issues – and on

boarders). E.g. http://www.nautilusminerals.com/irm/content/overview.aspx?RID=252&RedirectCount=1 (accessed 24.04.2018).

[4] http://www.mdpi.com/journal/minerals/special_issues/marine_minerals_gas (accessed 01.10.2018).

[5] The benthic zone is the ecological region at the lowest level of a body of water. It comprises the bottom, such as the ocean floor, the sediment surface and some sub-surface layers.

hydrothermal vents that are no longer active, in the sense that there is no geothermal activity going on (inactive hydrothermal vents);

2. cobalt-rich ferromanganese crusts on seamounts, which are underwater mountains of volcanic origin; and
3. polymetallic or manganese nodules[6] on plains on the deep ocean floor, called abyssal plains.

The impact of mining differs depending on deposit type as well as the geomorphological setting,[7] physical conditions, the scale of operations and the technology used for extraction. However, for the three types of mineral deposit mentioned above, the major impacts are similar:

1. loss of substrate,[8]
2. effects of mining on the seabed, the operational plume[9] (from sea bed extraction activities) and re-sedimentation[10] and
3. discharge plume[11] (from vertical transport and surface operations) and its effects on pelagic and or benthic fauna depending on the depth of discharge (ECORYS 2014, 24).

According to the same report, one of the most important effects is the actual removal of the minerals:

[6] 'Polymetallic nodules, also called manganese nodules, are rock concretions formed of concentric layers of iron and manganese hydroxides around a core. The core may be microscopically small and is sometimes completely transformed into manganese minerals by crystallization' (https://www.isa.org.jm/files/documents/EN/Brochures/ENG7.pdf (accessed 28.09.2018)).
[7] Marine geomorphological processes are processes associated with the action of waves, marine currents and seepage of fluids through the seafloor.
[8] The surface or material on or from which an organism lives, grows, or obtains its nourishment.
[9] Plume on the seabed generated by extraction of minerals.
[10] The process of moving sediment from one location to another.
[11] The discharge of waste water and fine particles generates a plume following initial on-board dewatering of the ore.

> Seafloor massive sulphides based in active hydrothermal vents (and the associated habitats) are expected to recover relatively quickly (months to years) while inactive sites will take considerably longer ranging from tens to hundreds of years. Nodule areas will likely take the longest time when it comes to recovery after the removal of the elements and may take tens to hundreds of years or even longer in heavily mined areas (nodule faunas may take millions of years to recover). Similarly crusts are expected to recover slowly meaning tens to hundreds of years (ECORYS 2014, 24-25).

For all three types of minerals, the sediment laden plumes near the seabed can spread kilometres beyond the mining sites and potentially 'smother seabed animals' (ECORYS 2014, 7). Sediment in the water column can decrease light penetration and temperature, which is likely to reduce plankton growth 'with knock-on impacts to the whole food chain' (ECORYS 2014, 25). In addition, ecosystems as a whole can be affected by the shift on sediment grain size.

Surface water pollution from ships and noise pollution from the vessels and underwater equipment may also have negative effects: 'With regard to noise pollution short-term masking effects on marine mammals are likely' (ECORYS 2014, 25). As with all mining activities, the disposal of tailings on land or at sea can have negative impacts.

A serious concern is that deep sea mining may lead to biodiversity loss or species extinction.[12] The discovery of hydrothermal vents being relatively recent, there is still considerable uncertainty about the degree of endemism of species to specific vent systems, and hence about the scale of the risk of species extinction. As a general observation, research suggests

[12] There is no single, universally accepted definition of neither 'biodiversity' nor 'species'. As Sandler (2012) argues, there are many legitimate conceptions of 'species', each being useful for specific purposes, for instance a specific explanatory purpose. In so far as the conceptions track real biological features of organisms, this does not imply relativism. In the context of the current dissertation, the 'form of life' conception proposed by Sandler (2012) (see also Crane and Sandler 2011), which understands species as 'groups of biologically related organisms that are distinguished from other groups of organisms by virtue of a shared *form of life*' (Sandler 2012, 6), seems adequate and reasonable. I will assume this conception, while recognizing that other conceptions may also be relevant. No substantial argument in this thesis depends on a particular definition of species.

It is common to measure 'biodiversity' or 'biological diversity' in terms of 'species richness' (see https://www.britannica.com/science/biodiversity (accessed 20.09.2018)). Species richness is a component of species diversity, and refers to the number of species within a biological community. Species diversity is determined not only by species richness but also by the relative abundance of individuals in that community (species abundance). See https://www.britannica.com/science/biogeographic-region/Components-of-species-diversity-species-richness-and-relative-abundance#ref588340 (accessed 20.09.2018).

that species richness is high at active hydrothermal vents, and lower at inactive vents. Active vents (also called active sulphides) host high-temperature (~350 °C) 'black smokers' that discharge metal-rich fluids. These metals precipitate at or below the seafloor to form polymetallic (especially copper and zinc) sulphides (Van Dover et al. 2018). In one article, hydrothermal vent ecosystems are described as 'natural wonders of the ocean':

> They exist as tiny islands in the unimaginably vast expanse of the deep sea; they are oases of vibrant and exotic life dependent on microbes that produce food using chemical energy through chemosynthesis. Biomass at active vents is dominated by species that rely on venting fluids and that can live nowhere else […] (Van Dover et al. 2018).

At active hydrothermal vents, which are quite rare (Van Dover et al. 2018) but relatively well studied (compared to inactive vents) (Gollner et al. 2017), mega- and macrofauna[13] 'show a high level of endemism in biogeographic provinces at both species (95%) and generic level (76%)' for fauna associated with polymetallic sulphides (Gollner et al. 2017, 78, cf. Moalic et al. 2011). The mega- and macrofauna are commonly endemic to the active vent environment (Tunnicliffe 1992).

Active vents inevitably become inactive ones (Zhang et al. 2016), which nonetheless can contain substantial mineral deposits (Boschen et al. 2013). Inactive hydrothermal vents provide potential habitats for microbial communities. Studies have shown that inactive hydrothermal vents harbour a completely different assemblage of microorganisms, compared to active hydrothermal vents (Zhang et al. 2016). The ('background') fauna at inactive hydrothermal vents 'resembles fauna of seamount communities with organisms typically being sessile, filter-feeding, long-lived and slow-growing, including taxa such as sponges, hydroids, corals, anemones, squat

[13] Deep sea macrofauna are benthic organisms or animals typically a few millimetres to several centimetres in size. Megafauna are animals > 50 mm. http://www.utas.edu.au/news/2017/11/28/467-explainer-what-are-marine-macrofauna/ (accessed 01.10.2018).

lobsters, ophiuroids and holothurians' (Boschen et al. 2013, see further Van Dover and Hessler 1990). Some macrofauna and nematode (worm-like organisms) species from active vent sites are also observed at inactive vent sites (Gollner, Miljutina, and Bright 2013, Levin et al. 2009).

Hydrothermal venting may be associated with seamounts, making them a target not only for crust mining but also for polymetallic sulphide mining (Boschen et al. 2016). Seamounts are considered to be 'hotspots' of species richness (Morato et al. 2010). However, knowledge is scarce about the seamount fauna in general, and about crust-associated fauna in particular (O'Hara and Tittensor 2010, Schlacher et al. 2014, Grigg et al. 1987).

Polymetallic nodules, also referred to as manganese nodules, occur in high abundance on the sediment-covered abyssal plains of all oceans, where sedimentation rates are low (Gollner et al. 2017). With regard to biological communities on nodules and in deep-sea sediments, species diversity can be high at both local and regional (spatial) scales. However: 'Remarkably little is known regarding the geographic distribution of the nodule associated fauna, such as whether species are unique to nodules or whether adjacent seamount crusts host these same species' (Gollner et al. 2017, 6).

In sum, there is substantial uncertainty about the fauna of deep sea environments eligible for mining activities. However, given the current state of knowledge about species richness and degree of endemism at all types of potential mining sites, and about the detrimental effects of some types of mining on vent habitats, the plausible scientific scenario is that species may go extinct as a consequence of deep sea mining operations in deep-sea hydrothermal vent environments, and that considerable loss of biodiversity and biomass may result from large scale deep sea mining.

Societal consequences of deep sea mining can be both positive and negative.[14] As pointed out in several reports (e.g. Earth Economics 2015, ECORYS 2014), there may be a rising need

[14] The distinction between environmental and societal impacts can be hard to maintain. For instance, we (presumably) tend to care most about environmental impacts that affect humans in some way, and thus think of environmental impacts in terms of societal impacts. However, it is clear that some impacts are best understood as

for minerals in the future, such as copper to meet demands for renewable energy sources. It is uncertain whether those needs can be met by recycling or other non-mining alternatives (ECORYS 2014), and some argue that deep sea mining may contribute to meeting at least portions of the demand (Earth Economics 2015). However, the significance of the contribution is debated (Van Dover et al. 2018).

Deep sea mining research activity may facilitate greater understanding of the deep-sea environment (ECORYS 2014). However, deep sea mining can also have negative effects on research activity and scientific progress in some areas, including biology and medicine (ECORYS 2014), which depends on keeping environments and biological material intact.

Social effects may depend on the context of particular projects. In connection with the Solwara 1 deep sea mining project in the Bismarck Sea outside of Papua New Guinea, Rosenbaum (2011) raises concerns about harm to local fisheries and possible social disruption (see also case study 2 in 'Precaution and Fairness: A Framework for Distributing Costs of Protection from Environmental Risks' in the current dissertation).[15] Moreover, as discussed in Study I of this dissertation – 'Existence Value, Preference Satisfaction, and the Ethics of Species Extinction'[16] – loss of species can affect human well-being in a negative way and may arguably be morally problematic regardless of its effects on human well-being.

Finally, deep sea mining may contribute to global climate change through pollution and massive energy use connected to operations on the deep sea (ECORYS 2014). Whether the overall climatic effects will be positive or negative depends on numerous factors, such as the temporal and spatial scale of deep sea mining activities, the number of operations that will take place, and whether they will replace or supplement terrestrial mining.

---

mere social impacts – for instance, the beneficial effects of copper supply on the economy – and that some impacts that we would call environmental have no (known) societal effects, for instance the extinction of some unknown species not necessary for any human needs or purposes (however, see the more complicated picture in the discussion of value concepts in Study I of this dissertation).

[15] Hereafter referred to as 'Precaution and Fairness'.

[16] Hereafter referred to as 'Existence Value'.

To sum up, deep sea mining involves novel technologies for obtaining minerals on the ocean floor in biologically and ecologically complex environments. It has not yet been carried out on a large commercial scale, but there are reasons to believe that it will be in the near future. Even though deep sea mining may involve important benefits for society, for instance through supply of important minerals and increased scientific understanding of deep sea environments, concerns have been raised about significant negative effects both on the environment and on society, notably in the form of species extinction, destruction of habitats, killing of sea-animals, pollution, social disruption, and lost opportunities for research and scientific progress. In spite of scientific uncertainty, we know that these are plausible effects. A different uncertainty concerns how such effects should be evaluated. That is a major topic of the current dissertation.

Risk, uncertainty, and the precautionary principle

I make frequent reference to the concepts of 'risk' and 'uncertainty' in the dissertation. In this section, I explain how the concepts are used. I also explain an important principle of environmental policy used in cases of risk and uncertainty – namely, the precautionary principle – and say something about how the dissertation relates to this principle.

In academic discussions, the concepts of 'risk' and 'uncertainty' are often understood in decision-theoretic terms. In decision theory, decisions under 'risk' denote situations where we have knowledge both of possible outcomes and their probability. Decisions under 'uncertainty', on the other hand, denote those situations where we have knowledge of possible outcomes, but *not* of their probability. Finally, decisions under 'ignorance' occur when we do not even know what the possible outcomes of our decisions are (Luce and Raiffa 1957, Peterson 2009).

The decision-theoretic framework is useful for distinguishing between different epistemic conditions for decision-making. However, it departs from common or everyday uses of the concepts of 'risk' and 'uncertainty' in several respects. For instance, in everyday language risk is understood to *involve* uncertainty, not serve as a contrast to it. Moreover, risk is commonly

associated with undesirability. For instance, when we say of someone that she is 'taking a risk', we mean that what she does may have undesirable consequences. Uncertainty, on the other hand, is used more widely to cover situations both of 'risk' and 'ignorance' (in the decision-theoretic senses).

The studies in this dissertation contain no technical, decision-theoretic discussions of risk and uncertainty, and no definitions from decision theory are presupposed. Instead, I use the concepts in a manner closer to everyday language. Unless something else is indicated, the term *risk* is used to denote either (a) an undesirable or unwanted event that may or may not occur – as in 'environmental risk' – or (b) the chance that some undesirable or unwanted event may occur – as in the 'risk of species extinction' (cf. Hansson 2018). (Which of the two senses the word is used in should at any time be clear from the context in which it is used.) The term *uncertainty* is used to denote either (a) an empirical or scientific lack of knowledge about consequences (which may include situations of 'risk', 'uncertainty' and 'ignorance' in the decision-theoretic senses), or (b) situations where we do not know or it is indeterminate how we should evaluate options, impacts, etc. The latter form of uncertainty is referred to as *evaluative uncertainty*.[17]

Some argue that decision-making in the face of environmental risk and uncertainty should – at least on some occasions – be guided by the *precautionary principle* (e.g. Steel 2015).[18] The precautionary principle comes in strong and weak versions. *Strong* versions state that some measure should be taken against serious environmental threats even if there is uncertainty about the nature of the threat, whereas *weak* versions state only that uncertainty should not be used as a reason for not taking action against such threats. An example of a strong version is the UNESCO definition: 'When human activities may lead to morally unacceptable harm that is

[17] Other forms of uncertainty include *normative uncertainty* (uncertainty about what should or ought to be done) and *moral* or *ethical uncertainty* (uncertainty about what the correct moral principles or ethical theories are). For further discussion of the decision-theoretic framework and the concept of uncertainty in the context of environmental policy, see Steel (2015).

[18] On the importance of the precautionary principle for environmental policy and law, see O'Riordan (1994), Trouwborst (2006), Steel (2015).

scientifically plausible but uncertain, actions shall be taken to avoid or diminish that harm'.[19] An example of a weak version can be found in the Rio Declaration of 1992: 'In order to protect the environment, the precautionary approach shall be widely applied by states according to their capabilities. Where there are threats of serious or irreversible damage, lack of full scientific certainty shall not be used as a reason for postponing cost-effective measures to prevent environmental degradation' (United Nations 1992).

When I speak of 'precautionary measures' in the dissertation, I mean any measure taken in order to reduce or avoid risk – from outright bans or moratoriums on certain activities or technologies, to less drastic measures such as requirements to do further research to map risks and benefits or to replace high-risk technologies with lower-risk technologies. By 'precautionary approach', I mean an approach involving precautionary measures – whether they are taken with reference to the (or a) precautionary principle or not.

I take no stand in this dissertation on whether any specific version of the precautionary principle should be invoked in the context of deep sea mining.[20] However, my studies have implications for the justification of precautionary measures. 'Hard Environmental Choices: Comparability, Justification, and the Argument from Moral Identity'[21] raises the issue of whether *moral identity* may influence our choice of approach to environmental risk in cases such as deep sea mining. Can statements concerning the kind of society we think we should be support a decision to take a precautionary approach to activities involving environmental risks, such as deep sea mining? 'Precaution and Fairness' discusses the justification of burdens imposed on stakeholders from taking precautionary measures. Distribution of burdens or costs can have implications for whether precautionary measures are implemented in a consistent manner – that is, in a manner avoiding the creation of risks that themselves call for precautionary measures.

[19] https://unesdoc.unesco.org/ark:/48223/pf0000139578 (accessed 19.03.2019).
[20] Concrete suggestions for precautionary management of deep sea mining are given by Durden et al. (2017). See also Mengerink et al. (2014), Halfar and Fujita (2002).
[21] Hereafter referred to as 'Hard Environmental Choices'.

This has to do with what is called the 'consistency constraint' on the precautionary principle (Steel 2015). Let PP be a version of the precautionary principle which implies precautionary measure X. PP is *inconsistent* if it would also imply that precautionary measures against X are called for, for instance because X would create a form of risk triggering PP. The *consistency constraint* says that a precautionary principle PP must, on pain of inconsistency, not imply a precautionary measure which itself triggers PP in this way. Arguably, unfair distribution of costs from taking precautions can involve unjustified harm (see 'Precaution and Fairness'). This implies that distribution of costs can involve a form of 'social risk', which can threaten the consistency of proposed measures (or more precisely: the consistency of their justification). Systematic consideration of distributional consequences can help reduce such risk.[22]

Environmental ethics

Environmental ethics emerged as an academic field in the 1970s as a response to urgent ethical challenges posed by growing pollution, global climate change, species extinction, environmental degradation and rapid technological development. At the dawn of this era, German philosopher Hans Jonas raised the question of how ethics could be reformed so as to be able to cope with issues of responsibility and uncertainty raised by the collective and global character of environmental problems (Jonas 1979).

Since then, environmental ethics has largely been preoccupied with the following broad philosophical issues: the value of nature; the moral status of non-human entities; responsibilities and obligations toward future generations; the relation between humans (culture/technology) and nature; the problems of risk, uncertainty, and collective responsibility.[23] The current dissertation

[22] For further discussion of the consistency constraint, see Sandin et al. (2002), Sunstein (2005b), Steel (2015), Kramer, Zaaijer, and Verweij (2017), Stabell (2017).
[23] Pioneering works are Sylvan (Routley) (1973), Jonas (1979), Leopold (1949), Rolston (1989), Callicott (1989), Næss (1990), and Taylor (2011). See also authoritative anthologies such as Light and Rolston (2002), and introductions such as Jamieson (2008). On the ethics of risk and uncertainty in the context of environmental problems, see, for instance, Munthe (2011), and Steel (2015). The question of anthropocentrism vs. non-anthropocentrism has moreover been central to environmental ethics. Aspects of this question are discussed in 'Existence Value'. A further

is most directly concerned with questions of value, risk and uncertainty as they relate to specific environmental issues relevant to the case of deep sea mining.

The summary at the end of this chapter provides more details on the studies and their particular topics. The point for now is that environmental ethics is not a unified theoretical or methodological framework, but a research programme focused around a set of central topics and problems. In brief, there are at least as many approaches to environmental ethics as there are approaches to ethics in general. In the next section, I will explain how my studies relate to the more general field of ethical theory and method.

A massive body of literature in environmental philosophy discusses whether, and in what way, non-human nature can be said to have 'intrinsic value'.[24] The concept of 'intrinsic value' can be understood in numerous ways – for instance, as non-instrumental, non-relational, or objective value (O'Neill 1992) – none of which are uncontroversial. I refer to the related concept of 'inherent value' in the studies. I take care to employ a minimally controversial interpretation: I take *inherent value* to mean the value that non-human nature, or at least living things in it, may have regardless of human preferences.

'Existence Value' discusses a preference-based value concept which has received much less attention from environmental philosophers, but which I argue deserves serious consideration in the environmental context. The concept of inherent value serves as a contrast to existence value in the study. In 'Hard Environmental Choices', inherent value is assumed to be part of the choice criteria used by the hypothetical ethics committee in the study. The arguments in these studies do not depend on any consensus about the concept or its exact nature. I make some suggestions in the Epilogue as to how inherent value can be further explored in the context of deep sea mining.

---

concept central to the discussions on environmental ethics is that of 'sustainability', which concerns (among other things) the interplay of socio-economic and environmental factors in economic development.

[24] Classic essays on the topic are collected in Light and Rolston (2002). Influential theories have been developed by Jonas (1979, 1966), Rolston (1989), Callicott (1989), and Taylor (2011).

Ethical theory and method

This dissertation offers discussions of substantive as well as more formal or procedural philosophical issues. 'Existence Value' examines a substantive issue of value. 'Hard Environmental Choices' and 'Precaution and Fairness' discuss procedures for choice and decision-making on the basis of multiple values and considerations. The studies were not written with a predefined methodological principle in mind, and contain theoretical and methodological explorations of their own. Nonetheless, some background assumptions about ethics unify the studies. In the current section, I explain the most important assumptions. On the basis of a distinction between two levels of ethical discourse – called 'factors' and 'foundations' – a framework is outlined for understanding the general approach taken to ethical issues in the dissertation.

Philosophical examination of real world issues requires close attention to facts about particular cases. This adds a layer of complexity to questions of theory and method. For instance, if an ethically ideal solution is unrealistic because of substantial conflicts of interest, then perhaps a 'non-ideal' approach will be appropriate (see 'Precaution and Fairness'). Or, if our theoretical considerations are strongly counterintuitive or in conflict with common-sense morality in particular cases, then this may – as Rawls (1971) among others suggests – indicate that we should bring our theoretical considerations into a form of 'reflective equilibrium' with the judgements or intuitions of common-sense morality. I try to give due consideration to the complexity of real life decision- and policy-making in the studies. The general approach outlined in this section should be regarded with this complexity in mind.

In the book *Normative Ethics*, Shelly Kagan (1998) introduces a distinction between *factors* and *foundations* in ethics.[25] In Kagan's terminology, factors or *normative factors* are those

[25] Another work operating with the distinction between factors and foundations is Donagan (1977). The distinction also has similarities with that made by Rawls (1971, 1985) between 'metaphysical' and 'political' moral theories – where the first kind is similar to foundational theory, while the second operates mainly on the factoral level.

considerations that are relevant for determining or deciding the moral status of an act.[26] Suppose we are to decide on the moral status of a particular act of promise-breaking. One normative factor would be whether the act had good consequences, or whether it promoted or violated important values, such as welfare or freedom/autonomy. Did it make the promisee better off than she would have been, had the promise been kept? Did it decrease her welfare or limit her autonomy? Other factors might include, for example, whether anyone was harmed by the act, whether the harm was intended or not, or what character traits (virtues) are expressed by the act.

Foundations, on the other hand, are theoretical considerations that provide explanations or justifications for what goes on at the factoral level: they explain why factors are relevant, in what way they are relevant, and how important they are in comparison to other factors. For example, utilitarianism of the maximizing sort purports to explain and justify why maximizing welfare is the ultimate normative factor, or the only factor with intrinsic value. In contrast, foundational theories of the Kantian (or deontological) breed purport to explain why following certain rules or principles is primary to considering consequences as a normative factor.

The distinction between factors and foundations can be hard to maintain in some cases. For instance, it may not always be clear what distinguishes accepting a form of consequentialism at the foundational level from employing a consequentialist principle at the factoral level. If I hold that the principle determining the importance of normative factors in a specific case is a certain consequentialist principle – for instance 'maximize total welfare' – is this not to say that the act in question is right if and only if it maximizes total welfare? In other words, is it not to allow a form of foundational consequentialism to determine the moral status of the act?

Not necessarily. In this example, someone accepting deontology at the foundational level could in principle allow a consequentialist principle to determine the moral status of the act, for

[26] The terms 'ethics' and 'morality' are treated synonymously in the dissertation, as is common practice in philosophy. This is the case also with concepts having the prefix 'moral' or 'ethical', such as 'moral issue' and 'ethical issue'.

instance if no significant deontological factors (such as doing harm) were involved.[27] Likewise, foundational consequentialism may imply deontology (or deontological constraints) at the factoral level. For example, accepting deontological constraints in the form of rules against harm-doing at the factoral level (i.e. in judging particular cases) is fully compatible with certain forms of foundational rule consequentialism.[28] Nonetheless, it should be conceded that it may not always be clear how to separate the two levels. Hence, the distinction between factors and foundations should be employed with some care. In the following, I use the distinction to explain in a rough manner the kind of assumptions I make about ethics in the dissertation.

I make some assumptions about normative factors. I assume that producing good results or outcomes is a relevant and important normative factor in deciding the moral status of acts. On this basis, I assume what can be called a 'minimalist consequentialist criterion'. This criterion says that, other things being equal, we should choose the (course of) action which has the best consequences.[29] This is a weak assumption. It says nothing about how we should evaluate the 'goodness' of consequences (to determine which are best). Moreover, it does not imply that we should evaluate actions on the basis of this criterion alone, or no matter what. It may be that other things are *not* equal – that is, there may be other factors or considerations that are (more) relevant for deciding what to do in a particular case or for evaluating our options.

Importantly, we often have insufficient knowledge about what the consequences of relevant courses of action will be. But we may have knowledge about other normative factors, such as that an action is one of harm-doing or violates human rights. In that case, we may have reason to choose an alternative course of action, which does not involve harm-doing or violation

[27] The distinction between consequentialist and deontological foundational theories is well established in the philosophical literature (see e.g. Rawls 1971, Kagan 1998). Classical works in foundational consequentialism are Bentham (1996 [1789]), Mill (2001 [1871]), Sidgwick (1981 [1907]). For a newer perspective, see e.g. Pettit (1997). The major classic in foundational deontology is Kant (1974 [1785/1786], 1974 [1788]). For an influential contemporary perspective in the form of contractarianism, see Scanlon (1998). Consequentialism can be thought to belong to a broader category of 'teleological' theories, which may also include virtue ethics (e.g. Kagan 1998).
[28] For further examples and discussion, see Kagan (1998, especially pp. 189 ff.).
[29] In cases of *ex ante* judgement, full knowledge of actual consequences is impossible and we must rely in some way on our knowledge of expected outcomes. Hence, I assume a probabilistic interpretation of the consequentialist criterion in the dissertation.

of rights, even though we do not know whether it is the best one with respect to overall consequences. In other cases, we may have sufficient empirical knowledge about consequences, but we are uncertain about how to evaluate them. That is, we are uncertain about what the *best* consequences will be or what kind of values or normative factors that are at stake.

It may be that the minimalist consequentialist criterion is best explained by a consequentialist theory. Or it may be better explained by some other, pluralistic theory. An example of a non-consequentialist pluralistic theory compatible with the criterion is the theory of David Ross (1930). Ross holds that we have a prima facie duty of 'beneficence' – i.e. to promote 'the good' in the sense of the welfare of others. However, he also holds that we have prima facie duties of justice, gratitude, and non-maleficence, which are irreducible to each other and to the duty of beneficence. The prima facie duties hold *ceteris paribus*. What our actual duty is in a particular case must be decided (if it can be decided) by considering relevant facts about the case at hand (see Ross 1930, 28 ff.). A further example of a non-consequentialist theory compatible with the minimalist consequentialist criterion could be a value-based virtue theory, where one value (among others) held by the virtuous person is the promotion of good results.

However, it is not the business of this dissertation to discuss what best explains or justifies the minimalist consequentialist criterion, and no assumptions are made about this foundational issue here. All that is assumed regarding the criterion, is that our evaluations and decisions should conform to it at the factoral level. But how, given the minimalist consequentialist criterion, can we decide the moral status of particular acts? I assume that the moral status of an act must somehow be a function of how different normative factors interact in particular cases. For example, the moral status of a particular act of promise-breaking seems determinable by a function of how producing or promoting good results interacts (*qua* normative factor) with other normative factors, such as avoidance of harm-doing; or how subfactors of promoting good results, such as promoting individual well-being, interact with other subfactors

such as promoting overall welfare, or with avoidance of harm-doing or subfactors of avoidance of harm-doing (such as harming as a means versus harming as an end).[30]

One way to explore what this function can be is to articulate and discuss *interaction principles*, understood as 'principles specifying how the various factors interact so as to determine the moral status of particular acts' (Kagan 1998, 183). Suppose that a member of a paramilitary group has placed a bomb in a densely populated area of a big city. Unless the paramilitarist is tortured, he will not reveal the location of the bomb. If we cannot find the bomb, at least five thousand civilians will die from the explosion. Suppose further that the normatively relevant factors are (1) the harm being done (to the tortured), (2) the (number of) civilian lives at risk, (3) the responsibility of the agent for the risk, (4) the availability of alternative means to remove the risk, and (5) the degree of certainty with which the act of harm-doing (torture) can be expected to save the lives at risk.

We now want to find a principle explaining how these five factors interact so as to determine the moral status of the act of harming (torturing) the para-militarist to make him give up the location of the bomb. A candidate interaction principle compatible with the minimalist consequentialist criterion in this case could be that 'when a thousand or more civilian lives are at stake, a certain extent of intentional harm-doing (torture) is justified when and only when (1) the one who is harmed is responsible for the threat, (2) harming this person is the only available means to saving the lives under threat, and (3) the lives will be saved with reasonable certainty'. Whether we should perform the act would then be a function of how the normative factors interact, given the interaction principle.

[30] I use the terms 'determine' and 'determinable' in a loose sense here. I do not assume that the moral status of an act can ever be fully determinate. However, I take it that, on the basis of interaction principles (see below) what we should do in some cases can be 'determinate' in the sense of 'quite clear' (or 'sufficiently clear', or something similar). For instance, if we think 'beneficiary pays' is a valid principle for distributing costs of precautionary measures in the environmental context, how costs should be distributed can be quite clear if the relevant facts about the case are known (see 'Precaution and Fairness').

In a philosophical discussion, an interaction principle such as this cannot simply be taken for granted. We have to discuss its justification. For instance, we can ask why the limit should be set at a thousand lives; why not a hundred, or ten? Or we can question whether it is *ever* permissible to inflict harm on individuals in order to save others from harm. And if we think it is, then how much harm can we inflict? Such questions require attention to ethical theory. For instance, to what extent should we let utilitarian considerations of quantities (of lives or pleasure/pain) be decisive? What can Kantian theories tell us about the specification of rules against harm-doing? Such theoretical considerations seem necessary in order to examine the justification of interaction principles at the factoral level.

The interaction of normative factors is explored in a variety of ways in the dissertation. The first study is mainly about analysing (preference-based) existence value as a potential normative factor in the environmental context. However, the study also discusses how this factor may interact with other relevant factors in the case of deep sea mining. I suggest that an approach based on the idea of prima facie duties may be reasonable in this case. 'Hard Environmental Choices' explores a case where relevant normative factors (criteria) are identified but fail to guide our choice. The study suggests that in at least some such challenging cases, the interaction of normative factors can be best understood in terms of 'deeper' questions of moral identity. 'Precaution and Fairness' suggests that examining how different principles of distribution interact in particular cases is crucial for arriving at fair distributions of the costs of taking precautions against environmental threats. Their interaction is examined with regard to theoretical considerations of justice and fairness.

No explicit reference is made to the concept of interaction principles in these discussions. Nonetheless, the concept captures a central idea at work in them – namely, that exploring the interaction of normative factors in particular cases is a key to understanding and resolving (at least some types of) moral issues. On this basis, the method of the current project can be described as being that of (1) identifying and analysing normative factors relevant to the case of

deep sea mining, and (2) discussing how relevant normative factors may interact in particular cases to determine the moral status of actions or courses of action that could be taken with regard to specific problems in deep sea mining (and similar activities), given the minimalist consequentialist criterion. This is not, however, meant to suggest a complete theoretical and methodological program. The studies in the dissertation offer further explorations of how ethical issues should be approached in complicated environmental cases such as that of deep sea mining.

Summary of studies

The dissertation contains three studies, corresponding to the three research questions raised at the beginning of this chapter. In the first study, 'Existence Value, Preference Satisfaction, and the Ethics of Species Extinction', a philosophical analysis is given of the concept of 'existence value', which refers to the value humans ascribe to the existence of something, regardless of whether it is or will be of any particular use to them. In the study, existence value is interpreted in terms of *preference satisfaction* and discussed with regard to the moral problem of anthropogenic species extinction.

Two main objections to using the concept of existence value in this context are addressed. The first is that on the preference satisfaction interpretation, the concept of existence value lacks moral importance, since satisfying people's preferences may involve no good or well-being for them. The second is that even if preference satisfaction can be linked to well-being, understanding existence value in terms of individual preference satisfaction is incoherent, since existence value reflects disinterested preferences that involve no benefits to the individual.

With regard to the first objection, it is argued that existence value can be based on a restricted version of the preference satisfaction theory, which is not vulnerable to the sceptical arguments about the link between preference satisfaction and well-being. With regard to the second objection, it is argued that the fact that existence value may involve disinterested preferences does not threaten the coherence or moral relevance of the concept, but suggests that

it does not fit smoothly into the 'utilitarian' or 'welfarist' framework that is commonly used to consider the concept. The final part of the paper suggests, on the background of a discussion of deep sea mining, some alternatives to standard utilitarian-style approaches for considering existence value in concrete cases involving a risk of species extinction.

While the main focus in the first study is on substantive questions of value, the second study, 'Hard Environmental Choices: Comparability, Justification, and the Argument from Moral Identity', discusses how choices can be made in cases where it is unclear what the best option is with regard to relevant values or choice criteria. In decision-making based on multiple criteria, situations may arise where agents find their options to be neither better than, worse than, or equal to each other. How, if at all, can a justified choice be made between such options? Are they incomparable?

Exploring a hypothetical case illustrating how this situation can arise in the environmental context – that of an ethics committee which is to make a choice between recommending or not recommending that a deep-sea mining project be allowed to proceed – this paper argues that the case is best understood as involving options that are comparable in the sense of being 'on a par'. On the background of a discussion of Ruth Chang's 'self-governance' theory of choice in cases of parity, it is suggested that in the environmental context, the idea of choices expressing a *moral identity* – reflected in statements about what kind of people or society we ideally think we should be – may lead us in the direction of a plausible solution to these hard cases.

Several authors argue for precautionary measures in the context of deep sea mining (e.g. Halfar and Fujita 2002). The third and final study of the dissertation, 'Precaution and Fairness: A Framework for Distributing Costs of Protection from Environmental Risks', tackles the question of how costs of precautionary measures can be fairly and effectively distributed. Daniel Steel and I address the issue by proposing a general framework for deciding how costs of precautions should be shared, which consists of a series of default principles that are triggered according to desert, rights, and ability to pay.

The framework is developed with close attention to the pragmatics of how distributions will affect actual behaviours. It is intended to help decision-makers think more systematically about distributional consequences of taking precautionary measures, thereby to improve decision-making. Two case studies – one involving a ban on turtle fishing in Costa Rica, and one involving a deep-sea mining project in Papua New Guinea – are presented to show how the framework can be applied.

Studies

I. Existence Value, Preference Satisfaction, and the Ethics of Species Extinction

Introduction

The current loss of biodiversity around the world is in large part due to human activity, including deforestation, urbanization, agriculture, and mineral exploitation. Extinction rates are currently 100–1,000 times higher than the background natural rate (Pimm et al. 2014, Ceballos et al. 2015). In this paper, the problem of anthropogenic species extinction is considered from an ethical perspective. The *value* that species are thought to have is crucial to their moral consideration. The question of the value of a species can be approached from an anthropocentric or non-anthropocentric perspective, i.e. we can ask about the value it has for humans, or the value it has independently of human desires or preferences. Moreover, it can be approached from a use or non-use perspective, i.e. we can ask whether it is useful or potentially useful for some particular human purpose, or whether it is valuable for some reason not related to use. In this paper, I consider the concept of *existence value*, which refers to the value humans ascribe to the existence of something, regardless of whether it is or will be of any particular use to them. The main research question is to what extent and under which conditions existence value should be used in approaching the moral problem of anthropogenic species extinction.[31]

Existence value is an important concept in the context of *ecosystem services*, which are of growing importance to environmental policy (Alcamo et al. 2003, Sukhdev et al. 2010, Barton et al. 2011). Although existence value has been extensively discussed, the concept has not received sufficient philosophical treatment. Methodological questions concerning how to determine and aggregate environmental values have been widely discussed in the literature on ecological

[31] Throughout the paper, 'species extinction' refers to 'anthropogenic species extinction'. Regarding the concepts of 'species' and 'biodiversity', see footnote 12 above.

economics (recently by Parks and Gowdy 2013, Kenter et al. 2015, and Kenter et al. 2016), as well as in related fields such as conservation biology (e.g. Justus et al. 2009). The role that environmental value concepts can play in policy making has also been discussed by several authors (e.g. O'Neill and Spash 2000). However, a systematic philosophical investigation of existence value is lacking in the literature. This gap in the literature is unfortunate, given the importance of existence (and non-use) value concepts in valuations of ecosystem services, and the serious ethical questions raised by such concepts. I endeavour to fill out parts of this gap in the current text.

Different conceptions of existence value can be found in the literature. My main focus will be on a 'preference-utilitarian' conception that appears in an influential United Nations Environment Programme (UNEP) framework for ecosystem services (Alcamo et al. 2003). This conception allows for a discussion of the role and importance of human *preferences* and *preference satisfaction* in moral deliberations about non-human nature, which is interesting from a philosophical and ethical perspective. Moreover, the dominating position the preference-based conception has in the discourse of ecosystem services (Alcamo et al. 2003) suggests that it should be carefully scrutinized.

I discuss two main objections to the preference-based concept of existence value. The first is that understood in terms of preference satisfaction, the concept lacks moral importance since there is no (morally significant) link between preference satisfaction and actual well-being. The second is that even were there to be such a link, it does not extend to existence value, insofar as existence value reflects disinterested preferences. With regard to the first objection, I argue that we should base the concept of existence value on a restricted version of the preference satisfaction view, which is not vulnerable to the sceptical arguments about the link between preference satisfaction and well-being, and which is compatible with other major views on well-being. With regard to the second objection, it is argued that the fact that existence value involves disinterested preferences does not threaten the coherence of the concept, but suggests that

existence value can be seen to involve different kinds of preferences, some of which do not fit into the 'utilitarian' or 'welfarist' framework that is commonly used to consider the concept.

The final part of the paper deals briefly with the question of the importance of existence value in concrete cases. The main purpose is to show that taking existence value into consideration does not require a utilitarian approach or adherence to 'welfarism', i.e. the view that individual welfare 'is the only thing with final or ultimate ethical value, the only state of affairs which we have moral reason to promote for its own sake' (Sumner 1996, viii). I suggest that an alternative approach might be to consider existence value within a pluralistic normative framework based on the concept of prima facie duties.

The discussion of the importance of existence value in concrete cases uses the example of *deep sea mining*. Deep sea mining involves retrieving minerals such as copper, gold, silver and zinc from the ocean floor at great depths in ecologically complex and biologically rich environments. Deep sea mining is relevant because it is considered to involve a risk of species extinction (Van Dover et al. 2017, ECORYS 2014). We have very little knowledge of many of the species thought to be at risk, and to that extent cannot say much about their actual or potential use-value. This makes the case an interesting testing ground for the concept of existence value.

A conceptual framework

An interesting conception of existence value can be found in the *Millennium Ecosystem Assessment* (MA) (Alcamo et al. 2003). MA was coordinated by the United Nations Environment Programme (UNEP), and is an important framework for considering *ecosystem services*, a concept of growing importance in environmental policy (Barton et al. 2011, Alcamo et al. 2003). The framework has formed the background for numerous environmental assessments and analyses. Recently, it was used as a basis of an environmental analysis by Earth Economics of what aspires to become the world's first deep sea mining project for copper minerals: the Solwara 1 project,

headed by the Canadian mining company Nautilus Minerals Inc. and located in the Bismarck Sea, off the coast of Papa New Guinea (Earth Economics 2015).

Ecosystem services are described in the MA as the products and benefits obtained from and provided by ecosystems (Alcamo et al. 2003). The MA groups ecosystem services into four broad categories according to how they benefit humans: 1) *provisioning goods and services*, which provide physical goods and materials, 2) *regulating services*, which provide regulation of e.g. climate and water quality, as well as keep disease organisms in check, 3) *supporting services*, which include primary productivity (natural plant growth) and nutrient cycling (e.g. carbon cycles), and are the basis of the vast majority of food webs and life on the planet, and 4) *cultural services*, which are functions that allow humans to interact meaningfully with nature (Alcamo et al. 2003).

Existence value is placed under the heading of *cultural services* – together with aesthetic, recreational and spiritual values, as well as medicinal, educational and scientific values. Existence value is distinguished from these other cultural values by being understood as a *non-use value*. While *use value* stems from the satisfaction or well-being that people might get – directly or indirectly – from the use of nature or natural resources (including such things as ecosystems and biodiversity), non-use value does not involve using something in nature for any particular purpose. Non-use value is not to be confused with 'option value', which according to the conceptual framework of the MA is the value we place on something that is not currently in use, but that might be used for some defined purpose in the future (such as biological material for medicinal use).[32]

The MA suggests that existence value is to be understood as a 'utilitarian' concept, based on 'a principle of human preference satisfaction (welfare)' (Alcamo et al. 2003, 19).[33] The idea

[32] In the MA, it is not entirely clear whether non-use value and existence value are used synonymously, or whether existence value is seen as a kind of use-value. However, it is common to distinguish three types of non-use value: altruistic value, bequest value and existence value (Kenter et al. 2015). My focus in this paper is on existence value only.

[33] In this paper, I focus on postive concepts such as well-being, desire, pleasure, etc., and put aside their negative counterparts, such as 'ill-being', aversion, pain, etc. These negative concepts obviously have moral relevance, but it is beyond the scope of the paper to deal adeqautely with both aspects.

seems to be that an agent considering something to have existence value gets 'utility'[34] in the form of preference satisfaction from knowing that it exists. No explanation is given in the MA of how the distinction between use value and non-use value is to be understood with respect to preference satisfaction. However, this can be clarified with reference to a distinction between intrinsic and extrinsic preferences.

Let me first say something about how I understand 'preference' and 'preference satisfaction' in this context. In the context of choice, 'preference' is commonly taken to refer to a stated or revealed ranking or rating. In a discussion of value and well-being, however, it is natural to understand preferences to refer (more substantively) to the wants, needs and desires of individuals (cf. Singer 2011). That is how I understand the concept of a preference here. Apart from the preference itself, a preference in this sense implies a preferring *subject*, which can be a human being or some other creature capable of having preferences, and a preferred *object*. The object of a preference can be either a thing, expressed in terms of *preference for*, or a situation (event, state of affairs, or the like), expressed as *preference that* (cf. Brülde 1998).[35] In the context of existence value, it is natural to presume that what is preferred is a certain situation, namely that some X exists.[36] On the situation interpretation, moreover, we can say that a preference that X is (roughly speaking) *satisfied* if and only if X actually holds.[37] Following philosophical convention, I

[34] 'Utility' in the preference-utilitarian scheme is commonly thought to be something that is 'set' by the agent, so that an agent's total utility function numerically maps out her preference ranking (Schick 1982). By contrast, in classical, hedonistic utilitarianism (in the style of Bentham), utilities were thought to be 'enjoyed' or 'pursued', for utility was simply pleasure, or 'that property in any object whereby it tends to produce benefit, advantage, pleasure, good, or happiness' (Bentham 1789, 8).

[35] Discussing the desire theory, Brülde (1998) argues that it is only the situation view (*desire that*) that has direct ethical relevance. The satisfaction of a *desire for* a thing – e.g. Janis Joplin's desire for a Mercedes Benz in the song of the same name – has ethical relevance only indirectly either because having the desire may be, if unfulfilled, unpleasant, or because it corresponds to a number of *desires that* (e.g. the possible desire of the Joplin-character that being in a situation where she is the owner of a Mercedes will give her a certain social status, etc.). On this background, Brülde argues, 'a plausible conception of well-being do [sic] not have to take our *desires for things* into account' (Brülde 1998, 169).

[36] Situations are complex and may involve several desires. For instance, my preference for the existence of a species, understood as a preference for a *situation* where the species exists (is protected, etc.), may involve a desire that the organisms will not suffer, that future generations may also enjoy the existence of the species, etc. The *preference* will then be for the situation 'as a whole', that is, as a 'compound' of the relevant desires.

[37] In other words, one should not confuse 'preference satisfaction' with 'satisfaction' in the psychological sense (e.g. with *feeling* satisfied).

will refer to preferences regarding particular situations as 'local', while preferences regarding situations in a broader sense, such as a life or part of a life, will be referred to as 'global' (Parfit 1984, Brülde 1998).

Extrinsic preferences are characterized by the subject or agent preferring something not for its own sake, but because of some 'extrinsic' reason – or, as von Wright puts it, characterized by the fact that 'a non-circular reason can be given for *why* x is preferred to y' (Von Wright 1972, 142). I can prefer a hammer over a rock because it is more useful for hammering nails. But in that case, I do not prefer the hammer for what it is in itself, i.e. not for its intrinsic properties. I prefer it for an 'extrinsic' reason, namely its usefulness. Intrinsic preferences, on the other hand – which are what mainly preoccupy us in this paper – involve preferring something for its own sake, and not because of its usefulness or relation to other preferences.

With this rough understanding of preferences and preference satisfaction in mind, let us now consider some basic challenges faced by existence value as conceived of in terms of human preference satisfaction. We will revisit the 'utilitarian' aspect in the final section.

The moral importance of preference satisfaction

The main objection to preference-based existence value in the moral context is that preferences and preference satisfaction lack moral importance.[38] What can it mean for something to have moral importance? As we have seen, existence value is connected to the idea of promoting human well-being (or welfare). The general idea that human well-being is good or has intrinsic value,[39] and that promoting it (or pursuing it, striving for it, or the like)[40] is morally important, is fundamental to much of the ethical literature in the western tradition (from Aristotle's

[38] Alternatively, moral standing or value.
[39] Alternatively, 'final value', in the sense that X is valuable as an end, or worth pursuing for its own sake.
[40] I will speak of 'promoting' for the sake of simplicity.

conception of *eudaimonia* to modern philosophers such as James Griffin and Derek Parfit).[41] I will take this general idea as my starting point. The assumption is that if preference satisfaction can be linked to well-being in the relevant way, this is an important step towards establishing the moral importance of (preference-based) existence value.

The preference (or desire) satisfaction theory is one of the major theories of well-being.[42] The core idea is that getting or achieving what we want is intrinsically important for our well-being, or for how well our lives go. This may be intuitively plausible. However, many authors argue that we should be suspicious of the idea that well-being consists in having preferences satisfied. A common line of argument is to point out that some preferences seem to be at odds with our self-interest or welfare. As Alexander Sarch puts it in a recent article:

> In some cases, our preferences are based on faulty information. In other cases, even when not misinformed, we prefer things that intuitively would not enhance our welfare. For instance, when our preferences are manipulated or otherwise not formed autonomously, their satisfaction does not always seem to enhance well-being […].[43] Moreover, even when our preferences are autonomous, we might still prefer things that intuitively do not make a positive contribution to welfare – as in certain cases of self-sacrifice or masochism (Sarch 2015, 142).[44]

These cases suggest that it is problematic to take the satisfaction of preferences as such to indicate well-being. However, this view is compatible with the view that well-being can consist (at least partly) in the satisfaction of preferences of a certain kind. For instance, we can say that only

[41] Aristotle (1976 [c. 400 BC]), Griffin (1986), Parfit (1984, 2011). *Eudaimonia* can be translated as 'human flourishing'. I do not believe it is a mistake to say that 'flourishing' involves a form of well-being in the broad sense of 'faring well', 'being well', etc.
[42] The major theories of well-being are (at least in the Anglophone philosophical literature) commonly understood to be (1) desire-fulfilment or preference satisfaction theories, (2) hedonistic theories, and (3) objective list theories (Parfit 1984, Griffin 1986, Sumner 1996, Brülde 1998).
[43] Here, Sarch refers to Sumner (1996). Cf. also Elster (1983).
[44] See further Brink (1989), Bradley (2007), Kraut (2013), Heathwood (2005). In the environmental context, see O'Neill (1993).

the satisfaction of what John Broome calls 'high-grade' preferences, i.e. preferences that pass some test of quality (Broome 1998, 272), should be taken to indicate well-being.[45]

To be sure, such a test of quality can be difficult to develop in a general manner. But it would presumably not be an impossible task to set up some criteria for existence value. Taking into consideration the problems cited above as well as the well-known problem of *morally* objectionable preferences (e.g. sadistic ones), they could *inter alia* include that preferences should be well-informed and not in conflict with common or widely accepted moral standards (e.g. should exclude sadistic preferences). Although it would be difficult to admit *only* high-grade preferences into our evaluations or deliberations on matters of species extinction/protection, it would presumably not be impossible to ensure that at least most preferences taken into account are approximately of this sort. Though not a perfect solution, it is preferable to excluding *all* consideration of preferences on the ground that some of them may not be high-grade.

However, some authors are not content with excluding preferences that are *objectionable* in some way from their conception of well-being, but hold the more radical position that there can be *no (non-contingent) relation at all* between preference satisfaction and well-being. In a recent article, Hausman and McPherson (2009) invoke Derek Parfit's argument against the so-called Unrestricted Desire-Fulfillment Theory of well-being (or prudential value) to show that the claim that well-being can consist in the satisfaction of preferences (understood in terms of desire fulfilment) is implausible. Parfit's argument proceeds by way of an example:

> Suppose that I meet a stranger who has what is believed to be a fatal disease. My sympathy is aroused, and I strongly want this stranger to be cured. We never meet again. Later, unknown to me, this stranger is cured. On the Unrestricted Desire-Fulfillment Theory, this event is good for

[45] Some argue that only the satisfaction of 'ideal' or 'hypothetical' preferences – for instance, preferences we would have if we were fully rational – can make our lives go better. However, I find it hard to believe that preferences we do not actually have can affect our lives in any significant way (cf. Griffin 1986). I therefore concentrate on the satisfaction of actual preferences as a candidate form of well-being.

> me, and makes my life go better. This is not plausible. We should reject this theory (Parfit 1984, 494).

Hausman and McPherson take Parfit's argument to suggest that there can be no more than a 'contingent relation' between preference satisfaction and well-being (Hausman and McPherson 2009, 10). It is important to notice, however, that Parfit's argument is directed against an *unrestricted* desire-fulfillment theory, which says that whenever a desire is fulfilled, this increases well-being. But in the context of existence value, there is no need to assume an unrestricted theory of preference or desire satisfaction. Existence value, as we have defined it above, is based on the presumption that the subject (S) knows or is aware that the object of her preference (X) exists. Hence, we need not assume that we can have preference satisfaction without knowing or being aware of it. All we need to claim is that insofar as (1) S prefers X, (2) X holds (the preference is satisfied in the 'situational' sense), and (3) S is aware that X holds[46] – this will involve a form of well-being for S (on the condition that the preference is high grade).

In other words, we can hold a *restricted* view that says that to the extent that our preferences are high grade, and as long as we are aware that they are satisfied when they are, well-being can consist in having such preferences satisfied. This view does not imply that well-being consists *only* or *always* in having preferences or desires satisfied. That is, the restricted view can be combined with the view that well-being can also consist in the fact that certain objective goods are obtained – i.e. with so-called *objective list* theories – or in pleasurable experiences, i.e. with *hedonistic* theories.[47] (On such a 'mixed' view, we could say that other things being equal, it is *better*

[46] Griffin (1986) discusses a similar restriction, which he calls the 'experience requirement' (p. 13). He rejects this requirement because he thinks it turns the desire theory into a mental state theory, which he rejects. However, this is arguably based on a confusion, see Sumner (1996, 127-128). One could specify the restriction further and say that the preference must be satisfied while we have it, i.e. it should be now-for-now, not now-for-then, etc. (Brülde 1998). When I speak of preferences in the following, I refer to actual now-for-now preferences.
[47] Cf. note 29 above about the major theories of well-being.

if these goods and experiences are also desired; and, perhaps, *best* if X is both pleasurable, desired, and on the positive objective list).[48]

To sum up, the discussion in this section suggests that the sceptic, insofar as she wants to deny that there is any (non-contingent) relation whatsoever between preference satisfaction and well-being, has to do a better job than merely providing a list of objectionable preferences and questioning the plausibility of the unrestricted desire-fulfillment theory (which I agree is implausible). If there are no better objections, we can safely hold on to the (intuitively plausible) view that there is an intrinsic connection between (high-grade) preference satisfaction and well-being (given certain restrictions).

Before we conclude about the moral importance of this link between preference satisfaction and well-being, a final objection should be considered. It could be held that even if there is such a connection between preference satisfaction and well-being, it is so 'minimal' as to have no important role to play in the moral context. In short, it could be maintained that only the satisfaction of preferences regarding 'basic needs' (such as security and subsistence) (Shue 1996), and perhaps possibilities (or structural 'capabilities') (Sen 1992) to for example move and speak freely, or for democratic participation, have moral standing. Such needs and capabilities, it can plausibly be argued, are more appropriately considered within a rights framework than a preference satisfaction framework. Therefore, there is no need to consider preferences.

Granted that basic needs and capabilities should be considered in terms of rights and duties, and that preferences have little relevance with regard to such needs and capabilities, it is problematic to say that no preferences other than those concerning basic needs or capabilities can have moral importance. The 'smaller things' in life arguably make up a substantial part of our well-being. Consider my (local) preference for Jim Jarmusch films over Hollywood blockbusters. On the basis of my arguments so far, we have reason to believe that getting this (presumably) high-grade preference satisfied is good for me. If we find the well-being of individuals morally

[48] For an example of a mixed theory, see Brülde (1998); cf. Parfit's 'compound view' in Parfit (1984).

important, it follows that the satisfaction of such preferences can have moral importance. Next, consider my (global) preference for leading a certain type of life, for instance the life of an academic philosopher. If we agree that high-grade preference satisfaction can have moral importance, I cannot see any reason why this sort of preference should not be granted such importance (under the right conditions).[49]

Assuming that human well-being is morally important and worth promoting, we can conclude that there are cases (other than those concerning basic needs and capabilities) where preference satisfaction (as a source of well-being) should be promoted. To the extent that existence value with regard to species extinction is such a case, there are reasons to promote it.[50]

Disinterested preferences

Even if we grant that high-grade preferences matter morally, we can still ask whether the kind of preferences involved in existence value are conducive to well-being. In the context of a critique of preference-based contingent valuation (CV) methods to measure existence value, Mark Sagoff claims that the view that existence value implies benefits to individuals in the form of preference satisfaction involves a 'logical problem' (Sagoff 2000, 1427). Allegedly, the problem arises because existence value implies 'disinterestedness', and thus 'may provide no direct or even indirect benefit to the individual: If the individual supports the policy option for disinterested reasons, the benefits CV is supposed to measure are not there' (Sagoff 2000, 1427).

Leaving aside for now the question of measurement (I will discuss it briefly in the final section), Sagoff's objection seems to assume that there is a contradiction between doing something for 'disinterested reasons' and getting benefits in the form of preference satisfaction

[49] About the ethical relevance of the desire to lead a life in accordance with a personal project, such as becoming a good lawyer or understanding Plato's theory of forms, see Williams (2006).
[50] Who should promote it? In the current context, the relevant duty-holders are mostly of a collective sort, such as public institutions or bodies/groups of policy-makers. I assume they have duties to promote welfare. Individuals may have duties to promote the preference satisfaction of other individuals in some cases, but I do not make any assumptions about that here.

from it. It is not entirely clear what Sagoff means by 'disinterested reasons' here. However, his objection seems to be a version of Sumner's disinterestedness-objection to desire theories, so I will take it to mean roughly what Sumner says – namely, that disinterested preferences involve no 'anticipated payoff for oneself' (Sumner 1996, 134).

Taken in this way, it is not obvious what the 'logical problem' is. Suppose that I prefer to support a policy of protecting an endangered species over another policy providing no such protection. That the species is protected can then *in fact* benefit me by satisfying my preference. But my preference for protecting it need not have been formed on the basis of self-interest. As Sumner himself points out, 'we are capable of finding enjoyment or satisfaction in doing good to others' (Sumner 1996, 135). In other words, we need to distinguish between the preferences we have, and our reasons or motivations for having those preferences. If I prefer that X, getting my preference satisfied can be good for me – even though my reason for preferring X is not that it will be good for me (to satisfy my preference) that X.

It may be (assuming that it makes sense to speak in this 'atomistic' way) that the gain in well-being I get from having disinterested preferences satisfied can be less than the 'loss' of well-being resulting from letting the interests of others come first – so that in the end I will be 'worse off'. But that does not show that there is a contradiction or a 'logical problem' involved. It only shows that there can be both gains and losses involved, and that the loss can be greater than the gain.

To the extent that Sagoff's claim is directed against the widespread assumption in cost-benefit analysis and neo-classical economics that all preferences are, ultimately, entirely self-regarding – i.e. that our motivation for ascribing existence value is solely the satisfaction of our own preference, not the welfare of others – it is fairly obvious that he has a point. It seems a mistake to assume that existence value is always and only a reflection of individual, self-regarding preferences. It is highly plausible that existence value can reflect concerns about the well-being of others, the justice of our actions, whether our actions and attitudes express certain virtues, and so

forth. The claim that our preferences are often other-regarding is moreover supported by empirical research (Irvine et al. 2016). Although this does not mean that existence value cannot also involve self-regarding preferences – in the sense that I prefer the existence of a species because the situation in which the species exists is best for me – it points to the need for methods that can take into account other-regarding aspects of existence value (Sagoff 1986, Kenter et al. 2015).

Finally, it can be pointed out that when we say that species have existence value, this may have little or nothing to do with the value we think the existence of the species has for humans. As Sagoff notes, one can prefer to protect the existence of a species not because it benefits ourselves or other people, but because we think it is worthy of moral consideration or respect in its own right (Sagoff 2000). This is (of course) plausible. But again, it does not imply that existence value cannot involve benefits to individuals in the form of preference satisfaction, as Sagoff seems to suggest. It only means that the preference may not be self-regarding (and this need not exclude individual benefit, as we have seen).

There are nevertheless good reasons to consider non-preference-based values in the context of species extinction. As has been argued for decades by environmental philosophers, making the value of non-human nature fully dependent upon human needs or preferences can be morally problematic. At least for individual non-human beings, it is plausible that they have a 'good of their own', and so have a form of objective or 'inherent' value, in the sense of a value independent of human preferences (Taylor 2011, cf. Hargrove 1992).[51] I will suggest some ways in which objective (non-preference-based) values can be considered in conjunction with subjective or preference-based values (such as existence value) in the final section.

[51] Some environmental philosophers, such as Holmes Rolston (1989), believe that a species can have a good of its own, and so a form of objective inherent value. For a critique of Rolston's view, see Sandler (2012).

Three kinds of preference

Based on the discussion so far, let me now try to systematize the different aspects of the concept of existence value. I have suggested that existence value can reflect self-regarding preferences. That is, I believe it makes sense to say that I prefer the existence of X because it will be better for me that X exists (and that it can in fact be good for me to have this preference satisfied). Moreover, the argument made by Sagoff and others that existence value may reflect social or shared values is convincing for reasons explained above. On this basis, I propose that the concept of existence value can be taken to involve (1) *self-regarding* preference (what I want to be the case for my own sake), and (2) *other-regarding* preference (what I want to be the case for the sake of society or other people).

As we have seen, preferences for preserving species regardless of the effect this has on human welfare can arguably be reflected in the concept of existence value. On this basis, I propose that existence value can also involve (3) *non-anthropocentric* preference – understood in the context of existence value as a preference one may have for the existence of something, regardless of its effects (or lack of effects) on human welfare or well-being.

Our other-regarding preferences can, I have argued, at least in an indirect way affect our well-being. The same goes for non-anthropocentric preferences. I may have preferences for the existence of species without any regard to human welfare, and it seems possible that satisfying such preferences can be at least indirectly (or derivatively) good for me. However, as noted above, satisfying such preferences may also involve a 'sacrifice' of well-being (Sumner 1996). Hence, there is a need to separate aspects of existence value that fit within a classical 'welfarist' framework, based on self-interest or self-regarding preferences, from those that do not, such as other-regarding and non-anthropocentric ones. Even if this may be difficult to do in a perfectly consistent manner, it should not be impossible to maintain the separation in a rough sense, which can be adequate for practical purposes.

In any case, the fact that having other-regarding and non-anthropocentric preferences satisfied may not be conducive to the subject's own well-being, and hence can imply challenges with regard to the measurement of existence value with methods such as contingent valuation, does not diminish the *moral* significance of the concept. That existence value may involve other-regarding and non-anthropocentric preferences rather suggests – unless one holds something like the radical view of 'ethical egoism' – a wider moral significance of the concept.

The importance of existence value

My main argument so far has been that (1) preference satisfaction can have moral importance, and (2) the preference satisfaction interpretation of existence value is coherent. However, I have not said anything specific about *how important* existence value is in the context of species extinction. How can we evaluate this? It can be difficult to do in a general manner. In the following, I will discuss it in relation to the case of *deep sea mining*. After pointing out some well-known problems with the standard 'utilitarian' approach, I will suggest some alternative directions evaluation can take in cases involving existence value, such as deep sea mining.

Defined as the area beyond the continental shelf, where water depths vary from 200 meters to 11,000 meters, the deep sea constitutes the world's largest biome and covers about 87 per cent of the ocean floor. The poorly documented deep-sea benthic ecosystems are thought to contain the deep ocean's largest reservoir of biomass and largest number of undiscovered species (UNEP 2007, TEEB 2012). Species residing in hydrothermal vent environments on the ocean floor are currently thought to be threatened by prospects of 'deep sea mining', which involves retrieving minerals such as copper, gold, silver and zinc from the ocean floor at great depths. Deep sea mining has at the time of writing this paper not been carried out on a commercial scale,

but many companies have received licenses to explore,[52] which raises concerns over the risk of species extinction (Van Dover et al. 2017).

An interesting fact about the deep sea mining case is that there is great uncertainty about the species in question and the habitats and ecosystems they are part of (ECORYS 2014). Some mining sites considered for deep sea mining are located at depths of several thousand meters. We know very little about the species that reside there – about the ecological and biological factors, but also, and to some extent related to that, about their use- and option-value. However, given that there are reasons to believe that some of them may be at risk of extinction from deep sea mining, it would be natural to ask whether we have reasons to protect them, even if they are of no direct use to us.

According to my arguments above, one reason for protecting them can be that people prefer that the deep sea species exist. How do we evaluate the strength of this reason? One approach, described in the MA as 'utilitarian', is to measure the preferences people have for the existence or protection of the species in monetary terms, by asking how much they are willing to pay to protect them. The importance of these preferences will then be determined by how they compare to other economic values. This approach may have the virtue of theoretical simplicity. However, as has been pointed out by a great number of authors, it is difficult to put the approach into practice in a sound manner, given (among other things) that different types of preferences and values are often at play that can be hard to measure on a single scale or criterion such as monetary value. While self-regarding preferences for material gains from mining may be measured in economic terms, it is doubtful (or at least highly uncertain) whether the other-regarding or non-anthropocentric preferences people may have for the existence of a deep sea species can be adequately measured by asking how much they are willing to pay to protect them (Kenter et al. 2016, Sagoff 1988).

[52] https://www.isa.org.jm/deep-seabed-minerals-contractors (accessed 20.11.2017)

Moreover, if we take a pluralist view and assume that there are different 'basic' values and considerations at play as well – such as subjective (preference-based) values on the one hand, and objective (in the sense of non-preference-based) values, such as inherent value, on the other – then the utilitarian approach, to the extent that it assumes the complete commensurability of values, seems even more problematic.

I will not take a stand on whether some form of utilitarian or welfarist approach may be appropriate after all. But if we do not want to make the controversial assumption of commensurability – that is, if we want to start out from a more pluralistic view – what can we do? If we want to remain in the vicinity of decision theory, an alternative could be to use techniques based on Multiple Criteria Decision Making (MCDM), where options are ranked on the basis of multiple scales or criteria to generate information about relative weights.[53] Alternatively, we can take a more substantive normative approach. Let me suggest in general terms how this can be done.

We can, for instance, start out from the idea that the different values and normative factors involved are grounds for prima facie duties of relevant actors. These duties hold other things equal, but can be overridden – although not cancelled out – by stronger duties in particular cases (e.g. Kagan 1998, cf. Ross 1930). Suppose we believe that public institutions have a prima facie duty to promote human well-being (including high-grade preferences). On the other hand, we also think they have a prima facie duty to protect nature (for its own sake). Against this background, existence value can be *one* factor in deciding what the actual or 'all things considered' duty of the relevant actors should be. Its importance as a factor will depend on the strength of the relevant preferences.[54]

---

[53] For state of the art surveys on MCDM, see Greco, Figueira, and Ehrgott (2005). It is disputed whether MCDM type approaches can completely avoid the incommensurability problem (e.g. Peterson 2007).

[54] The measurement of strength does not require a cardinal scale, such as monetary value. Preferences can be measured on an ordinal scale – for example, a scale from 'very weak' to 'very strong'. I leave methodological issues regarding the measurement of preferences aside here; but I assume that it will be possible to identify preferences, and that they can be (in some way) measured adequately for at least some practical purposes. Methodological issues regarding the identification and measurement of preferences remains to be clarified.

For example, suppose that a certain deep sea mining project can provide society with substantial amounts of copper, which is important for the development of renewable energy sources. At the same time, the project threatens a deep sea species. The species is considered to have negligible use- and option-value, which implies that from a *use* perspective, we have no (or – factoring in inevitable uncertainty – at least very weak) reasons to protect it. On the other hand, let us presume, the 'good' or the inherent value of the (members of the) species is a *pro tanto* reason to protect it. It seems that in this case, we have reasons both to mine and not to mine, and it may not be clear what we should do. In this situation, if people's preferences regarding the existence of the species are strong, then this consideration could provide a stronger reason not to mine. If preferences are weak, then existence value provides a weaker reason not to mine.[55]

Of course, this general approach needs to be fleshed out further. The point is that there are viable alternatives to monistic, utilitarian-style approaches for considering existence value in particular cases. The sketched 'prima facie' approach shows that one does not have to be a preference-utilitarian or a proponent of welfarism to take existence value seriously in the context of species extinction. We may instead, for example, be pluralists or pragmatists who recognize the value of human preference satisfaction in the moral context, but do not think that environmental issues can or should be solved by purely utilitarian or economic means.

Conclusion

This paper began by asking to what extent and under which conditions existence value should be used in approaching the moral problem of anthropogenic species extinction. I have argued that existence value can involve benefits for human individuals in the form of preference satisfaction. Assuming that human well-being should be promoted, this suggests that existence value should be taken into account in evaluating activities that come with a risk of species extinction.

[55] The question of how we should proceed if it is unclear how options compare 'all things considered' is addressed in Study II of the dissertation.

I proposed that existence value can be understood to involve self-regarding as well as other-regarding and non-anthropocentric preferences. This suggests that the concept has moral significance beyond concerns for individual human well-being. It also suggests that the concept is ill-suited for consideration within a purely welfarist or utilitarian framework based on self-interested preferences. Using deep sea mining as an example, I sketched an alternative approach for considering existence value in concrete cases, based on the notion of prima facie duties. The approach can accommodate 'objective' values, such as inherent value, which may be an advantage from the perspective of environmental ethics.

II. Hard Environmental Choices: Comparability, Justification, and the Argument from Moral Identity

Introduction

In environmental decision-making, decision procedures that rely on a single scale or criterion of measurement, such as cost-benefit analysis, are attractive since they offer the prospect of a clear ranking of options – even a cardinal one if based on monetary values – which can provide a firm foundation for rational decisions.[56] If building a dam has the value of one million dollars, while protecting the ecosystem threatened by the dam has only half the dollar value – and the sole criterion is monetary value – then clearly we should build the dam!

However, the assumption that diverse values can meaningfully be measured on a single scale or criterion is problematic given what can be called the *incommensurability principle*, which states (roughly) that there is an irreducible plurality of values such that those values cannot be 'commensurated' or reduced to a single, common measure. For instance, the aesthetic value of a river landscape can be held to be incommensurable with economic value, in the sense that you cannot measure the former in terms of the latter.

The incommensurability principle is philosophically contested (e.g. Regan 1997). Taking a stand on this issue would require a thorough examination of value pluralism, which is beyond the scope of the current paper.[57] The starting point of my discussion is that at least from the practical perspective of providing decision aid to actors in concrete cases, adopting the incommensurability principle in some form seems reasonable and perhaps even necessary. The reason is that whether decision aid is provided for a single actor or takes place in a multi-actor

[56] Cost-benefit analysis is widely applied in environmental decision-making (see Sunstein 2005a, Steel 2015).
[57] For classic expositions of incommensurability and its connection to value pluralism, see Williams (1981), Nagel (1979), Berlin (1990, 2002), Stocker (1990).

setting, it is rare that actors have in mind a single, well-defined criterion. More often, it will be necessary in a decision process to take into consideration various points of view dealing with, for instance, environmental aspects, financial aspects, human resources, and security (Köksalan, Wallenius, and Zionts 2013).

The current paper discusses a problem that may arise in decision-making based on multiple considerations or criteria.[58] In a well-known example, Joseph Raz (1986) asks us to imagine a person facing a choice between a career as a lawyer and one as a clarinettist. Suppose that the relevant choice criteria for this person are economic security and musical interest. In this case, the person may find no career better than the other with respect to *both* economic security and musical interest – while not being indifferent between the two careers either. Such cases can arise in the ethical domain as well. In this paper I examine how, in the environmental context, choices can be made between options that are considered neither better than, worse than, or equal to each other with respect to the relevant criteria.

I discuss a hypothetical (but realistic) case illustrating how this choice situation can arise in the environmental context – namely, that of an ethics committee struggling to decide whether to recommend that a certain mining project be allowed to proceed or not, since they do not find any option clearly better or worse than the other, but do not find them equally good either. It is natural to presume that a justified choice between options requires that they can be compared in some way. But what is comparability, and why is it important for justified choice?

I begin by considering the so-called 'trichotomy thesis', which says that there are only three comparative relations in the context of evaluation: 'better than', 'worse than', and 'equal to'. The thesis implies that a case like that of the ethics committee, where none of these relations hold between options, is a case of *incomparability*. If no comparison can be made between options

[58] There exist a host of more technical (formal and quantitative) literature on what is called Multiple Criteria Decision Making (or Multiple Criteria Decision Analysis). See, for instance, Greco, Figueira, and Ehrgott (2005). The problems discussed in the current paper can presumably arise in most forms of decision-making based on multiple criteria.

based on the three standard relations, can agents still make a rationally justified choice between the options? I discuss four proposals for how this can be done – 'maximilizing', 'satisficing', 'absolutizing' and 'secondary considerations' – and argue that they are all problematic in the environmental context. I then discuss the view that the trichotomy thesis is false, since options that are not comparable in terms of the three standard relations can nevertheless be compared in terms of 'parity' (Chang 2002) or 'rough equality' (Parfit 1984, Griffin 1986, 1997). I argue that the ethics committee case is best understood as involving comparability in the sense of parity or rough equality.

Against this background, I go on to assess a proposal by Ruth Chang (2013, 2009b, 2009a) for how choices can be made in cases of parity. Arguing that Chang's 'self-governance' approach is problematic in the ethical context because of its reliance on voluntarism, I suggest a related approach based on the Aristotelean idea of *moral identity*, which concerns the kind of society we think we should be in light of some conception of the good life. I argue that while the self-governance approach implies an arbitrariness that is problematic for the type of environmental choices discussed in the paper, considering choices in light of how they affect and are affected by moral identities may provide a substantive, non-arbitrary basis for justification in cases such as that of the ethics committee.

The ethics committee

Suppose that an ethical advisory board or ethics committee is appointed by an official body of policy-makers to advise it on whether to allow a mining company to proceed with a copper mining project on the ocean floor.[59] The mandate of the committee is to clarify ethical issues and,

[59] The example is based on the real case of so-called deep sea mining, which involves retrieving minerals such as copper, gold, silver and zinc from the ocean floor at great depths. Several companies and states are involved in exploration. At least 28 contracts have been given by the International Seabed Authority (ISA): https://www.isa.org.jm/deep-seabed-minerals-contractors (accessed 01.03.2018). For a pioneer project within exclusive economic zones, see: http://www.nautilusminerals.com/irm/content/overview.aspx?RID=252&RedirectCount=1 (accessed 24.04.2018).

if possible, to make clear recommendations to the decision-makers as to whether the project should proceed or not. The proposed mining site is located in a biologically rich and ecologically complex environment.

Suppose further that the ethics committee is in possession of substantial knowledge about the economic costs and benefits of the project, as well as the expected environmental impacts. Careful economic analyses have been carried out which conclude that the benefits of the project for society, notably in the form of minerals needed to make the 'green shift' to renewable energy sources, significantly exceed the economic costs. On the other hand, the expected damage to the environment is significant. The hydrothermal vent environment, where the mining is to take place, may be irreparably damaged, individual benthic organisms will be harmed and/or killed in the mining process, and significant species or biodiversity loss may be unavoidable.[60]

Not wanting to assume commensurability, the committee employs a multiple criteria approach to the problem. To simplify, say that the committee considers the ethical legitimacy of the project on the basis of two ethical dimensions. On one dimension, environmental aspects are considered, including animal welfare and a concept of the inherent (or non-anthropocentric) value of nature.[61] On the other, human welfare is considered on the basis of economic and other anthropocentric values. No lexical priority is assumed to exist between the dimensions, and no dimension is taken to be reducible to the other.

Considering animal welfare and the inherent value of nature, the committee thinks that damage to the environment should be given most weight. This favours the no-mining option (for instance by way of a precautionary approach). However, concerns of human welfare, notably represented by the cost-benefit analysis, favour the mining option. Taking both dimensions into consideration – i.e. 'all things considered' – the committee finds no option to be clearly

[60] On environmental risks of deep sea mining, see e.g. ECORYS (2014), Van Dover et al. (2017). On the state of scientific knowledge, see ECORYS (2014).
[61] The concept of the 'inherent value' of nature is meant to capture the view or intuition many people have that (at least some things in) nature has moral standing independent of human preferences, needs, etc. Classic philosophical essays on this topic are collected in Light and Rolston (2002).

preferable to the other.[62] Still, they are not indifferent between the options: it is not the case that they do not care which alternative is chosen, or find the alternatives exactly equally good (or bad) with respect to their criteria.

If the committee wants to avoid suspending judgement on the matter, what can they do? Can they make a rationally justified choice between the alternatives? A natural place to start in answering this question is to ask whether the committee can in any way compare the options in order to evaluate their relative merits. What kind of comparability, if any, exists between options in cases where agents do not consider any option better or worse than any other, while they are not considered equally good either? If there is no comparability between the options, can they still make a justified choice? Different theories have been developed in the philosophical literature to answer these questions. I will discuss the most important theories in the following sections, using the ethics committee as a 'test case' to see how they hold up in the environmental context.

Incomparability and weak comparability

Some authors argue that choices where options are considered neither better than, worse than, or equal to each other, are (at least in practice) cases of *incomparability* (Raz 1986, Anderson 1997). We can call this position 'incomparabilist'. Incomparabilists assume what has been called the 'trichotomy thesis' (Chang 2014), which holds that the logical space of evaluative comparisons is limited to three comparative relations: 'better than', 'worse than', and 'equal to'. If none of these relations hold between options, it is claimed, the options are incomparable, since in such cases there are no comparative terms available to assess the relative merits of the options.

[62] In philosophical discussions, 'preference' is often treated as synonymous with 'want' or 'desire' (cf. Study I). This is not how the concept is used in this paper. In accordance with standard terminology in decision and choice theory, I take 'preference' to refer to a ranking of options in terms of a value relation, such as 'better than' or 'worse than'. I make no metaethical assumptions about the ultimate 'source' of such ranking here. That said, I do find it natural to think that the form of ranking relevant in the ethics committee case reflects a *value judgement*, which expresses a belief on the part of actors about what is good, best, etc. – as opposed to mere desires or wants of the committee.

It can be held that rational choice in cases where none of the standard comparative relations hold – I will refer to them as 'hard choices' (Chang 2012) from here on – can be based on 'weak comparability', which does not require complete overall preferences over options (Martinez-Alier, Munda, and O'Neill 1998).[63] The gist of the idea is that there can be comparability on some dimensions or criteria without there being a more comprehensive or 'strong' comparability (such as commensurability) between the alternatives (O'Neill 1993).[64] For instance, even if agents' overall preferences between apples and pears are incomplete, they can still prefer apples to pears on one criterion, say price, while pears are preferred to apples with regard to taste. However, it is not clear what implications weak comparability has for choice in cases such as that of the ethics committee, where a comprehensive or all-things-considered judgement about what to do is sought. Weak comparability does not, it seems, exclude incomparability on this 'higher' level. What does this imply for the possibility of justified choice?

If agents cannot compare their options on the basis of the three standard comparative relations, this bars the form of justification most prominent in economics and decision theory – namely 'optimizing' – since optimizing requires that we choose an alternative that is *at least as good as* the others. Does this mean that no rationally justified choice can be made between them? It may be that even if we cannot make a meaningful evaluative comparison of our options, we may still be able to justify our choice of one option over another. At first glance, at least three alternative forms of justification seem available in cases of incomparability (understood in terms of the trichotomy thesis). Let us consider their plausibility in the environmental context.

[63] In decision theory, an agent's preferences are said to be 'complete' if the agent either has a strict preference for some item in the choice, or is indifferent between them. Otherwise, they are 'incomplete'.
[64] It is not entirely clear whether 'weak comparability' is suggested by O'Neill as a normative thesis about justification, or merely a descriptive thesis about a certain type of choice situation. From the discussion in (O'Neill 1993), both interpretations seem possible. I interpret it here as a normative thesis, or at least a thesis supposed to have normative implications.

Three strategies of justification

The strategies can be called 'absolutizing', 'maximilizing' and 'satisficing' (Chang 2014). Let us consider them in turn. In the case of absolutizing, justification proceeds on the grounds of some absolute principle or rule. No matter how *x* compares to *y*, if choosing *x* is required or permitted by the principle or rule in question, we are *ipso facto* justified in choosing *x*. For instance, the act of fulfilling a promise can be justified 'absolutely' on the grounds of a principle stating that we should always fulfil our promises.

Even though absolutizing might be a plausible form of justification in some cases, it seems problematic as a strategy for environmental decisions. Absolutizing economic growth, for instance, would mean that we could do whatever we wanted to the environment as long as economic growth was secured. Absolutizing species protection would imply that we could not extinguish harmful bacteria such as those leading to smallpox. It is unlikely that decisions such as these could be justified without any reference to the merits or demerits of alternative options. Rather than being viewed as absolute considerations, economic benefits and environmental damage are considerations that must be weighed against each other when in conflict.

The notion of 'maximilizing' (Sen 2017, cf. Raz 1997) provides an alternative strategy that is arguably open to both incomparabilists and comparativists (the latter being those who believe that comparability is required for justified choice) (Hsieh 2007). In contrast to optimizing, which says that to choose *x* over *y* is justified if *x* is at least as good *y,* maximilizing requires only that we choose an alternative that is *not worse than* any other alternative. In choosing between items that are not better than, worse than or equal to each other, we are justified in choosing any alternative, since none is worse than the others.

Even if we overlook the theoretical complications that have been pointed out with regard to maximilizing (Chang 2014), the maximilizing solution is problematic on account of the phenomenology of the type of choices we discuss here. Consider again the example of the ethics committee. The question posed to them is 'should we allow the mining project to continue or

protect the environment by not allowing mining?' Suppose their answer is: 'We can choose either alternative, because neither is worse than the other'. This answer does not seem satisfactory, and I believe this can be explained by the fact that in the type of case we are discussing, the ethical stakes are very high: If the choice is between preserving a unique and valuable ecosystem and gaining some (sufficiently substantial) material benefit to society by destroying it, we presumably need to know more in order to be confident of our decision than that neither option is worse than the other.

The same holds for so-called 'satisficing', i.e. for justifying a choice by the fact that the option one chooses is 'good enough'. To the extent that it can work as a coherent form of justification at all,[65] being 'good enough' may not provide strong enough justification in the environmental context, where choices involves high ethical stakes. In Chang's terminology, the *justificatory force* is lacking that is necessary to determine the choice (Chang 2014).

In sum, even though absolutizing, maximilizing and satisficing may in principle provide rational justification for choice in cases where none of the standard comparative relations hold between items, they are all problematic for hard choices in the environmental context. While absolutizing can be said to have a justificatory force that is so strong that it becomes implausible in concrete cases, the problem with the remaining two strategies, maximilizing and satisficing, is that they cannot provide *enough* justificatory force in the type of cases we discuss.

Secondary considerations

Hsieh (2007) proposes the following solution to the problem of underdetermination of choice due to lack of justificatory force: If the fact that *x* is not worse than *y* and vice versa with respect to choice value or covering consideration *V* does not provide us with the justificatory force

[65] Chang (2014) argues that it cannot.

necessary to determine choice, a 'secondary consideration' – not itself part of V – might be introduced that may provide such force.

For instance, when a male applicant for a job is neither better than, worse than nor equal to the female applicant with respect to relevant qualifications, then 'gender' might be introduced as a 'tiebreaker'. In this situation, gender is not part of the choice value 'relevantly qualified'; it is irrelevant to qualification and does not make one candidate better than the other. At the same time, if neither candidate is worse or better than the other, then considerations of gender can justify choosing one over the other.

Could the ethics committee reason in a similar way? What secondary consideration could the committee appeal to that would not itself be part of the choice value but that would still be important to the choice? The secondary consideration could be some kind of non-moral or pragmatic value. Presumably, it must be a significant value, not something like 'pleasantness of the breakfast menu for mining employees'. Could we find such a value in the case at hand?

It may be that the ethics committee could appeal to pragmatic considerations such as whether the time would be right to allow the mining project, with respect to public opinion. If the decision they recommended would be sufficiently unpopular, the decision could threaten the *ordre public*. It seems unlikely, however, that a decision regarding the case at hand could threaten the *ordre public*. In fact, I suspect that in most cases, a tiebreaker performing a similarly significant role to that of gender in Hsieh's example will be hard to find. But the solution should not be dismissed out of hand, and can be kept in mind as we go on to consider some further suggestions.

Parity and rough equality

Since the incomparabilist strategies considered above all seem to fail or to be of limited use, let us now look at some alternative views. As will be recalled, the incomparabilist view relies on the trichotomy thesis, which says that there are only three possible comparative relations. That thesis

has been contested. Notably, Ruth Chang (1997, 2002, 2012, 2014) argues that there is at least one further comparative relation, which she calls 'parity'. Parfit (1984) and Griffin (1986) argue that even though, strictly speaking, the logic of comparisons is restricted by the three standard relations, items that are neither better than, worse than or equal to each other can be 'roughly equal' and in that sense comparable.[66]

The argument behind these claims is that in many cases where none of the standard comparative relations hold between items, a closer look will reveal that they are nevertheless comparable in the sense that we can make a positive evaluative judgement about their relative merits – which intuitively is what evaluative comparisons are all about.[67] To find out if alternatives are comparable in this sense, Chang (1997) proposes a 'notable-nominal test'. To see how the test works, let us apply it to the choice faced by the ethics committee.

The first step in the test is to find a 'nominal' version of one of the options. A nominal version of the mining option could be 'environmentally harmful mining with only trivial benefits as a result', or 'with benefits only for a small group of already wealthy shareholders'. Intuitively, this option – we can call it 'pointless mining', as opposed to the original option which we can call 'mining' – is clearly worse with respect to ethical considerations (representing the all-things-considered judgement in this case, since the two ethical dimensions or set of ethical criteria are part of or constitute the field of 'ethical considerations') than the option 'not mine and protect the environment', or 'no mining' for short. In other words, 'pointless mining' is comparable to 'no mining' with respect to ethical considerations.

The next step in the test is to gradually improve the 'pointless mining' option. By a series of small improvements in resulting benefits for society, we can bring 'pointless mining' closer and closer to 'mining'. At some point, they will be (roughly) identical with respect to benefits for society. Chang's point is that it is implausible that by any such small improvement, the options

[66] Parfit uses the terms 'rough comparability' (Parfit 1984) and 'imprecise comparability' (Parfit 2011).
[67] In contrast to positive judgements, negative judgements are for instance the judgement that *x* is *not* better than *y*, or *not* worse than *y*.

will change overall from being comparable to being incomparable. Hence, the 'mining' option must be comparable to the 'no mining' option.

It is contested whether the notable-nominal test – also known as the Chaining Argument (CA) – proves the existence of a whole new comparative relation, as Chang claims (e.g. Boot 2009). But even if understood as a borderline case or a case of vagueness, CA could nevertheless be taken to imply parity in the sense of rough equality between items (Griffin 1986, Parfit 1984).[68] Important in our context is that CA strongly indicates that even if none of the standard comparative terms can be used to compare items, a positive evaluative judgement can in some cases, such as that of the ethics committee, be given on their relative merits in terms of rough equality or parity. (I do not consider parity and rough equality to be significantly different concepts in the practical setting. I will therefore treat them as equivalent in the remaining discussion, and use the term 'parity' for simplicity.)

The crucial question remains: What are the implications of parity for justification in hard choices? Can parity help us make a comprehensive judgement and avoid the triviality of weak comparability? One obvious negative implication is that, since parity cannot tell us whether any option is at least as good as any other, optimization is excluded. However, Chang thinks parity can also have positive implications for practical reasoning in hard cases:

> Parity expands the range of cases in which justified choice is possible; choices between items about which practical reason might otherwise appear to be silent are in fact choices between comparable items and thus within the scope of practical reason. Parity, it might be said, is what gives practical reason a 'voice' in hard cases' (Chang 2005, 333).

[68] CA may appear to be an instance of the 'sorites paradox' or a problem of vagueness in our language or concepts (cf. Broome 1997). Chang (2014, 2002) argues against this, but the matter is largely unsettled. If the problem is vagueness, then it does not follow from CA that the trichotomy thesis is false. In any case, this does not exclude 'rough equality'.

It is not clear how the claim about giving a 'voice' to practical reason should be understood. However, it seems clear that saying that items are on a par is to say something positive about the value relation between the items. Intuitively, this seems to matter in the context of practical reasoning. But *how* does it matter, exactly?

It seems to follow from the fact that no alternative is better or worse than the others in cases of parity that we have sufficient reason to choose either alternative and are therefore rationally permitted to do so. In turn, this implies that we are rationally permitted to choose arbitrarily between them, for instance by flipping a coin. This implication leaves parity in tension with the phenomenology of ethical choices such as that faced by the hypothetical ethics committee: that the committee is not indifferent between the alternatives suggests that they do not experience the choice as one where they can choose arbitrarily between the alternatives.

As mentioned above in connection with maximilizing and satisficing, a plausible explanation for the lack of indifference in these cases is that ethical choices are experienced as too *important* to be decided arbitrarily; they *matter* too much to leave the decision to chance or whim. How then, if at all, can a non-arbitrary justified choice be made between alternatives in hard choices, understood as cases of parity? To answer this question, it is necessary to examine the nature of the *reasons* involved in these cases. That is the starting point of the proposal to which we will now turn.

Self-governance

Chang (2009b, 2013) suggests that in cases where one fails to have more, less, or equal reason to choose one alternative over another – or where we have reason to choose one alternative over another, but how much more reason is indeterminate – we are justified in *creating* a reason through an act of willing. The reason thus created is called a 'voluntarist reason', as opposed to what are called 'given reasons', which (for simplicity) can be understood here as our choice criteria.

Chang's proposal is called 'hybrid voluntarism', since it does not rely exclusively on voluntarist reasons, but also on reasons 'given' in other ways, such as by desires or normative facts (Chang 2013). In a choice situation characterized by incomplete preferences, hybrid voluntarism would suggest that completeness can be obtained by the agents *deciding* by an act of will to have a preference for one option over the other, even though they have no reasons for their preference on the basis of their choice criteria. 'Very roughly, when you will something to be a reason, you *put yourself* behind some consideration that, as a logical matter, counts in favor of one of the alternatives' (Chang 2013, 180).

To explain how this might work, Chang introduces the notion of *rational identity*, which is 'a description of your normatively ideal self – a loosely unified way of understanding the reasons that justify doing what you have most reason to do' (Chang 2009b, 261). The basic idea is that we create our rational identities through the choices we make. Therefore, when alternatives are on a par, we should ask: What kind of rational identity do we create (or reinforce) if we choose *x* over *y*? The answer can be used to put ourselves behind a particular option, the choice of which confirms and in a sense constitutes who we are (as rational beings). Chang calls the procedure *self-governance*.

The idea of self-governance is appealing in our context, since presumably we want our choices regarding environmental issues to be rational, in the sense of being responsive to what we have sufficient or decisive reasons to do or not to do – and to express as much as possible an ideally rational way of being. The main problem with Chang's proposal is its commitment to voluntarism. The notion that we create reasons and identities through willing certain options is problematic for at least two reasons. Firstly, it seems obscure when appealed to in choices of a collective nature. Say that the ethics committee gave the following answer to the question of whether they would recommend the 'mining' (a) or 'no mining' (b) option:

> We find no decisive given reasons to choose either *a* or *b*. Neither are they exactly equal in value. They seem to be on a par. Given this situation, we collectively *will* that the promotion of human welfare is significant for us, and so we give ourselves an additional reason to recommend *a* over *b*.

Suppose further that, in order to clarify this reasoning, someone commented: 'Ok, so what you mean is that you *want* welfare promotion to be a weightier reason, and thus advise us to choose *a* over *b*'? And the committee answers: 'No, far from it. We do not *want* welfare promotion to be a weightier reason. We *will* it; we used our autonomy to create a weightier reason for *a*'.

A legitimate response to this would be: 'But *why* do you "will" it?' I do not think that the committee can simply answer that it is because they created a reason for it, even if they add that the process behind it is one of 'autonomous willing'. In the case of individual choice, such as that of choosing a career, the will or desire of the agent may have some authority in justifying the choice. In collective decisions, such as that of the ethics committee, it is not clear why the will of decision-makers should have this kind of authority – or, if it is a collective will that decides, how this collective (autonomous) willing is supposed to work.

Secondly, if I can will reasons, then it seems that I can will *any* kind of reason. To use the example by Cohen (1996), if I am a Mafioso, I can will a reason to bust the kneecaps of someone owing me money. This is problematic even if this willed reason does not outweigh my reasons not to do it. Moreover, it is unclear whether Chang suggests that willed reasons can be created in all situations but only have normative significance for our choice in situations where our options are on a par, or that they can only be created when the options are on a par. The first suggestion seems implausible because of examples such as that of the Mafioso, the second because the restriction seems arbitrary: if we have the ability to create reasons, it is hard to see why it would be restricted to cases of parity.

As I have argued in this paper, we should avoid arbitrariness in ethical hard choices. In the following section I propose that thinking about the role of identity in a more ethically

substantive way – namely as involving questions of *moral identity* – can help us deal with the problem of arbitrariness, and moreover suggests an account of justification from identity that is more appropriate to collective decisions.

Moral identity

We can understand moral identity as concerning who we – as individuals or collectives – ideally think we should be in light of some substantive ethical view, for instance on what constitutes a good life.[69] Collectively, moral identity can be understood along two dimensions. The first concerns the extent to which an identity is shared by the community or society in question. We can call this the communal dimension. The second concerns the history of this community – decisions made by the community in the past, their legal and political practices, institutions, and so forth. We can call this the historical dimension.

As an example, consider a moral identity which we can call the 'green' identity. The green identity, let us suppose, is guided by a conception of the good life called 'sustainable development'. Since the concept of sustainable development – defined early on as 'development that meets the needs of the present without compromising the ability of future generations to meet their own needs' (World Commission on Environment and Development 1987) – arguably should include some conception of non-human welfare and the inherent (non-anthropocentric) value of nature, let us say that sustainable development is comprised of a concern for present and future human needs in addition to these non-anthropocentric values.

The green identity connected to this vision of sustainable development might, in its collective form, be captured in statements such as, 'we are a society that cares for present and future human and animal welfare and respects the inherent value of nature'. As we shall see, invoking such identity statements is different from merely invoking the reasons for them, since

[69] The classical theorist of moral identity is Aristotle (1976 [c. 400 BC]). I develop my view along broadly Aristotelean lines. For alternative views, see e.g. Taylor (1989) and Korsgaard (1996).

moral identity may also influence the importance and weight given to reasons. The gist of the proposal I want to sketch here is that choices can be made in hard cases on the basis of reflections on moral identity. Let us examine the rationale of this proposal.

When agents face choices between 'notable' and 'nominal' options – for instance a mining project with only trivial benefits for humans and very serious impacts on the environment, versus a very effective protection plan – questions of *identity* seem of little relevance. There is, it seems, no need in such cases to consider complicated questions of who we are, what kind of society we should be, and so on. The choice situation considered in this paper, however, is not one of notable and nominal options, but one of parity (or, if one does not accept the argument against the trichotomy thesis, one of 'incomparability' or weak comparability) with respect to the criteria used or the normative factors involved. In this situation, I suggest, it is natural to ask questions about moral identity.

In the ethics committee case, such questions can address how the relevant society identifies itself. It may align more strongly with a green moral identity, expressed in statements such as, 'We are a society that needs clearly outweighing reasons to support a mining project involving a risk of species extinction and serious environmental degradation', perhaps favouring a precautionary approach to deep sea mining, or alternatively with a statement such as, 'We are a society that puts maximization of human well-being first in decisions having to do with the environment', perhaps favouring a cost-benefit approach.

Before considering the proposal in more detail, let us look at it in light of the concept of rational choice. Chang's introduction of voluntarist reasons to resolve hard cases can be interpreted as an attempt to give a more substantive account of rationality than that suggested by given reasons. When the 'thin' rationality of given reasons fail, then other types of reasons can justifiably be introduced, implying a 'thicker' or more substantive conception of rationality and rational choice. As we have seen, however, voluntarist reasons of the kind that Chang considers are problematic. What kind of reasons are suggested by the concept of moral identity?

Korsgaard (1996) argues that without letting ourselves be guided by a conception of our identity, we shall 'have no reason to act and to live' (123). This may be an overstatement. However, moral identity seems fundamental at least in the sense that the idea we have of ourselves as moral agents influences the reasons we are disposed to see as significant, or should be disposed to see as significant for our choices. This indicates that moral identity is distinct from an abstract value to be promoted or honoured, or with respect to which comparisons can be made. Rather, moral identity can be seen as (constituted by) a set of traits or dispositions which makes us approach matters in certain ways, including a disposition to take certain values, options or forms of comparison into consideration in the first place.[70]

So conceived, one could raise the objection that it is hard to see what kind of reasons moral identity statements can provide. In fact, one could say, the approach appears no less arbitrary than the creation of voluntarist reasons. It can be pointed out in response that moral identity can be non-arbitrarily invoked on the basis of the communal and historical dimensions explained above. It is not the case that the ethics committee could choose to invoke *any* moral identity statement. If the statement lacks strength along the communal and historical dimensions, and no other convincing reasons are given for invoking it, then invoking it becomes arbitrary and for that reason should not be thought to provide a reason for a decision.

Particular identities can be morally problematic, however, even if they are strong on these dimensions. Consider the moral identity of a community of slave owners in the 18th century American South. Could they not appeal to the strength of their identity as a 'white' community in considering whether they are justified in keeping non-whites as slaves? It may be implausible that ethical considerations are on a par in this case and hence that such an appeal would be justified. Nevertheless, the example shows that our choice (or cultivation or endorsement) of moral identities can have serious consequences. This may be the case even if at some point, and with

[70] For a closer examination of how character traits, or virtues and vices, may influence our choices and actions in the environmental context, see Sandler (2007).

regard to some type of situations, endorsing the identity was justified. Suppose that the ancestors of the community of slave owners justifiably endorsed an identity of strong 'in-group' ties in the face of outside threats to their existence. As this identity became more deeply rooted in the practices of the community, it could influence later generations' evaluation of the practice of slavery: 'The slave is not one of us, so he does not deserve equal moral consideration'. In this way, a moral identity that had some justification in earlier times, can become problematic as times change.[71]

This shows that we should be careful how we choose or interpret our reasons for decisions that are on a par, as the reasons we give for such choices are likely to travel to non-par situations; we should be careful not to endorse moral identities that are likely to lead to harmful consequences when extended more broadly. In the deep sea mining case, for instance, endorsing the 'cost-benefit' identity in a situation of parity could influence choices we later make in situations of nominal-notable options, because the reasons making an option nominal may be given less weight than they would have been given had a green identity previously been endorsed (in a strong enough manner to influence choice). And this applies similarly for the green identity, of course.

Crucially, the fact that identities can be morally problematic suggests that in addition to constraints given by the historical and communal dimensions, moral identities are subject to substantive normative constraints. That some moral identities are more problematic than others, implies that we can have *good reasons* to choose or endorse certain identities over others. While Chang's self-governance approach implies that we can decide to 'put ourselves behind' any reason given in the choice, then, the moral identity approach suggests two types of constraint on our preferences and choices in cases of parity: one is 'contingent', in the sense that it depends on historical and sociological facts about particular societies; the other is normative, in that it

[71] This can be related to Charles Taylor's idea of changing moral frameworks in Taylor (1989).

depends on the reasons we have for preferring one identity over another (independently of contingent facts about our societies).

Several justifiable moral identities may be at play in a pluralistic society. It is not the case, however, that *any* identity will be justifiable: we have no reason to endorse a moral identity with bad influence on our choices. Moreover, the reasons provided by considerations of moral identity need not be implied by the initial choice criteria. The green identity may dispose us to be more respectful in our dealings with the natural environment. Even if this general attitude of respect for nature is not itself part of the choice criteria, it can affect the weight we give to a criterion such as 'avoidance of harm to the hydrothermal vent environment'. Endorsing the green identity may in this way provide the ethics committee with a stronger reason not to recommend the mining project, or to recommend a precautionary approach to mining.

Conclusion

I have argued that the ethics committee case should be understood as one where options are on a par. I suggested that in the environmental context, where decisions are ethical and collective, a choice between options on a par can be justified on the basis of moral identity statements – i.e. statements about what kind of society we think we should be, in light of some conception of the good life.

Some issues regarding the moral identity proposal have not been sufficiently dealt with. One concerns how, exactly, the historical and communal dimensions of moral identity should be understood. Another is the worry that it will not always be clear or agreed upon when options are on a par, and thus when the moral identity strategy is justified. Finally, the question of how we should deal with competing moral identities in a pluralistic society deserves closer examination. These complexities point to the need for further discussion of the proposal.

What seems clear from the arguments in this paper is that finding rationally and morally defensible solutions in hard environmental cases can require reflection on deep and difficult

questions concerning who we are and who we think we should to be – that is, of collective moral identity. This result has significant implications for our choice of approach in these cases: it is plausible that questions of moral identity do not have 'expert' answers, and that broad public debate and deliberative practices of a collective kind are required in order to arrive at adequate statements.

III. Precaution and Fairness: A Framework for Distributing Costs of Protection from Environmental Risks

1. Introduction

This paper explores an issue that has not been adequately examined in the literature on the precautionary principle and environmental ethics more generally: how to fairly distribute the costs of taking precautionary measures against risk. Precautionary measures can come with substantial costs for the parties involved, and this fact has implications for how precautions should be implemented. For example, carbon taxes are one precautionary measure against anthropogenic climate change and have been introduced in several jurisdictions, including the Canadian province of British Columbia. However, since carbon taxes can impose a heavy burden on low-income people, households in British Columbia that have an income below a certain level are compensated through a tax credit system.[72]

The question of how to fairly distribute costs of environmental precautions is significant. Unfair distributions of costs and benefits are morally problematic, and precautions that distribute costs unfairly may encounter resistance because they are viewed as illegitimate. Yet the question of fair distribution of costs of precautions has received relatively little attention in literature on the precautionary principle and environmental ethics. Some have argued that distributional consequences should be taken into account when applying the precautionary principle (Dickson 2005), and others have championed one distributional principle or a set of principles in specific contexts, such as climate change (e.g. Neumayer 2000, Caney 2005, Page 2008). However, these proposals are importantly incomplete. Recognizing the importance of fairly distributing costs of precautions raises the question of how to distinguish fair from unfair distributions. And while a

[72] http://www2.gov.bc.ca/gov/content/taxes/income-taxes/personal/credits/climate-action (accessed 25.04.2017).

particular distributional principle or set of principles may be salient in one context, a general account of the topic requires a framework for considering several principles in tandem. Indeed, such a framework is needed even for the analysis of single cases, wherein multiple plausible but competing principles may be invoked, as is illustrated by cases we examine in Section 4. In this paper, therefore, we identify principles relevant to fairly distributing costs of precautions and propose a framework for how to jointly apply them in a variety of circumstances.

We begin by articulating distributional principles relevant to our context and by examining their rationale in different kinds of considerations of responsibility and justice, notably desert, rights, welfare, and equality. On this basis, a framework is proposed for the fair distribution of precautionary costs. The framework consists of a default principle, called Risk-Initiator Pays, in addition to further principles that can be invoked when there are strong moral reasons for sharing burdens. The structure and dynamics of the framework are illustrated in Figure 1. The framework is intended to help decision-makers think more systematically about distributional consequences of taking precautionary measures, thereby to improve decision-making. Two cases – one about a ban on turtle fishing in Costa Rica, and one about a deep-sea mining project in Papua New Guinea – are presented to show how the framework can be applied.

The term 'precautionary measure' (for short, 'precaution') is used in this paper to refer to any measure taken against a risk or hazard in order to reduce or negate it. Such measures can be everything from outright bans or moratoriums on certain activities or technologies, to less drastic measures to control or reduce risk, such as requirements to do further research to map risks and benefits or to replace high-risk technologies with lower-risk technologies. In many cases, such measures involve a reference to or application of the precautionary principle,[73] which is highly influential in environmental policy (Trouwborst 2006, O'Riordan 1994, Steel 2015). But as we

[73] About the precautionary principle, see the section 'Risk, uncertainty and the precautionary principle' in the Introduction.

conceive of them in this paper, precautions may or may not involve an explicit use or reference to the (or a) precautionary principle.

2. Conflicting principles for distributing costs of precautions: a 'non-ideal' approach

In *The Idea of Justice*, Amartya Sen (2011) asks us to imagine three children who disagree about who should get to play with a flute. Anna says she should have it because only she knows how to play it; Bob says he should have it because he has no other toys; and Carla says she should have it because she is the one who made it. All three agree on the facts, but they disagree about who should get the flute because each prioritizes a distinct principle of justice: libertarian right to the fruits of one's labour, which favours Carla, or economic equality, which favours Bob, or hedonistic utilitarianism, which (arguably[74]) favours Anna. The point of the story is that focusing on a single principle is inadequate. Justice, Sen insists, requires balancing reasonable principles or conceptions of fairness that often conflict in concrete cases. Such an approach to justice is unlikely to take the form of a universal theory that provides a transcendental conception of the perfectly just world – and indeed it need not do so.

A more promising approach is to develop proposals that guide comparative judgments about more or less just social arrangements in a specific type of context. Moreover, in addition to principles of justice, such proposals should also pay attention to the pragmatics of how proposed reforms would affect actual behaviours. Sen's approach to justice, then, is an example of 'non-ideal theory' (Valentini 2012), and the framework we propose here with respect to fair distribution of costs of precautions is advanced in this spirit.[75] Our framework integrates several

[74] As Sen (2011) observes, utilitarianism could also favor Bob on the basis of decreasing marginal utility, or Carla on the grounds that the right to keep what one has produced encourages economic productivity (13-14).
[75] While non-ideal theory seeks to give due consideration to people's actual behavioral patterns in moral and social situations – for instance the risk of partial compliance to agreements – ideal theory assumes an 'ideal' social world consisting for instance of rational moral agents tending to act in full compliance to agreements. For what is often considered a prime example of ideal theory, see Rawls (1971).

potentially conflicting principles relevant to fair distribution of costs of precautions in order to guide comparative judgments about justice, while giving pragmatic concerns about incentives their due consideration.

The first step in developing such an approach is to identify relevant principles and to explain how their potential for conflict raises difficult questions of justice. That is what we do in this section. The principles (Table 1) have been selected and refined primarily on the basis of two criteria. First, we aimed to include principles that have been discussed or applied in connection with the distribution of costs of precaution or in related contexts such as public finance and climate justice. Second, we aimed for a set of principles that were comprehensive insofar as making it possible to consider the responsibilities and rights of all of the actors affected by precautions or their costs. Achieving the second of these aims required modifying some principles to make them more general. In addition, to avoid ambiguity we divide the notion that beneficiaries may have responsibility to bear the costs of precautions into two separate principles, one concerning beneficiaries of the risk-generating activity and the other pertaining to beneficiaries of precautions. To our knowledge, the latter of these principles has not been previously discussed in the literature.

We begin by considering the Polluter Pays Principle (PPP). This is a principle for the allocation of costs from preventive pollution control (Gaines 1991, OECD 1972), and a means to internalize the costs of pollution, insofar as pollution constitutes what economists call a negative externality. In its general form, uncoupled from the specific problem of pollution, PPP can be taken to state the quite intuitive notion that the one who harms the environment or public health, or stands in danger of doing so, should bear the costs of compensating for or reducing that harm.[76] Translated into the language of risk, PPP can be said to state the likewise intuitive notion that whoever initiates the risk should bear the costs of precautions. Since polluters are not the only relevant agents, we will call this principle Risk-Initiator Pays (RIP).

[76] Cf. also the 'contribution to problem' principle in the literature on climate justice (e.g. Neumayer 2000; Page 2008).

Principle	Abbre-viation	Explanation
Risk-Initiator Pays	RIP	Those who initiate the activity that generates the risk should pay in proportion to their contribution to the risk
Ability to Pay	ATP	Those who are most able to afford the costs should pay
Beneficiary of Activity Pays	BAP	Those who benefit from the activity that generates the risk should pay
Beneficiary of Precaution Pays	BPP	Those who benefit from taking precautions against the activity that generates the risk should pay

Table 1. Four distributional principles used in our framework. A catch-all category, called 'Others Pay', is added to the framework in Section 3 (see Figure 1).

A rationale for RIP is that the causal connection between the agent (the risk initiator) and the action (the risk imposed) gives a reason to place the burden of taking precautions on the one causing the risk rather than on someone who is not causally responsible. A further rationale can be found in desert-based views on just distribution. One prominent group of desert-based views emphasizes the connection between desert and contribution (Miller 1976, Riley 1989). Such views may be taken to imply that the costs of taking precautions should be distributed in a way that does not place greater burdens on people than they deserve for having contributed to the risk (Miller 2009).

The general plausibility of the desert view is perhaps best seen by looking at the implications of denying the moral importance of desert to just distributions. Consider the claim that in the case of two people performing a job, it is of no moral importance whether the one who contributes most to getting the job done receives at least an equal salary to the one who contributes less. Even if welfare for some reason would be maximized by giving a greater salary

to the one who contributes less, it may still seem unfair to do so. A plausible reason for this is that the person who contributes most does not deserve to get paid less than the one who contributes least. On the contrary, it can be argued that the person who contributes most deserves a greater salary – at least if she has also put a greater effort into getting the job done (Sadurski 1985, Milne 1986).

Similarly, desert can be a plausible ground for holding that someone contributing negatively to the welfare of others, for instance by putting them at risk of serious harm, have a greater obligation to reduce the risk or compensate for the harm than those contributing less to the risk. For example, it is natural to suppose that countries that have historically contributed greater amounts of greenhouse gas emissions to the atmosphere have a proportionately greater obligation to bear the costs of climate change mitigation.[77] In some cases, desert should arguably be overridden by other concerns, such as ability to pay (see below). What is important for now is that desert should be one consideration among others in distributing the costs of precautions, and that it can provide a rationale for RIP in particular cases.

A second principle relevant to our context states that the burden of taking precautionary measures 'must be placed on those most able to afford it' (Thompson and Kennedy 1996). It resembles the principle of public finance called 'ability to pay', which states that those who have the means should share more of the burden of public services. Moreover, versions of this principle have been central to discussions about how costs related to combatting climate change can be fairly distributed (e.g. Shue 1999, Neumayer 2000, Caney 2005, Page 2008). In concordance with standard terminology in these discussions, we call this principle Ability to Pay (ATP).

A central rationale behind ATP is that it is unfair to impose costs on poor individuals or groups that are not able to afford them. In that light, ATP reflects considerations of justice

[77] For discussions of 'historical responsibility' for climate change, see e.g. Shue (1999), Neumayer (2000), Caney (2005), Page (2008).

directed at protecting the least advantaged. Several considerations of this sort can be found in the literature. One prominent example is John Rawls's 'difference principle', which allows for unequal distributions as long as they benefit the least advantaged. This principle is fair, Rawls argues, because it is what reasonable people would choose for their social institutions in a hypothetical 'original position', where a 'veil of ignorance' ensures that decision makers have no knowledge, at the time of deciding, of their own chances of finding themselves in the worst position (Rawls 1971).

Also, some welfare-based views, notably utilitarianism, stress maximizing welfare for those least well off based on the thesis that each unit of a thing maximized will be marginally less valuable the more one has of this thing (diminishing marginal utility). Hence, the negative value of costs from taking precautions will be marginally less for those more able to afford it. However, utilitarianism will only protect the least advantaged so long as this maximizes overall utility. Distributing costs according to ATP could further be supported by the 'priority view', which states that benefitting people matters more the worse off these people are. On this basis, it is argued that we should prioritize the worst off even when this does not maximize overall well-being (e.g. Arneson 2000). Finally, sufficientarians argue that threshold values of welfare or 'contentment' exist that no individual should fall below (Frankfurt 1987).[78] This supports avoiding imposing costs that would push individuals or groups below some acceptable threshold of welfare.

All these views have been heavily debated, and none of them prove that ATP should be an overriding principle of fair distributions. Fortunately, that is not what we are after. What is significant is that taken together they give substantial theoretical support to the weaker but, for our purposes, sufficiently strong claim that the concern for the worst off expressed by ATP should be, at the very least, a serious moral consideration when distributing the costs of precautions.

[78] Cf. also the 'capabilities approach' as discussed for instance in Nussbaum and Sen (1993).

In some cases, RIP and ATP may coincide. For example, when considering how the costs of climate change mitigation should be distributed on a global scale, it might be argued that wealthier industrialized nations are both the primary risk initiators as well as the most able to pay. However, RIP and ATP pull in opposite directions when those who initiate the risk are less well off. Such tensions are illustrated by the cases discussed in Section 4. Consequently, an adequate framework for the just distribution of costs of precautions must be able to address examples in which RIP and ATP suggest conflicting recommendations.

A third principle suggests that beneficiaries should bear costs of precautions (Goodin 2013, Goodin and Barry 2014, Lawford-Smith 2014). One recent interpretation of this principle in the context of climate justice states that 'being an innocent beneficiary of significant harms inflicted by others may be sufficient to ground special duties to address the hardships suffered by the victims, at least when it is impossible to extract compensation from those who perpetrated the harm' (Barry and Kirby 2017, 285). In order to separate beneficiaries of the activity from beneficiaries of precautions (see below), we suggest a principle called Beneficiary of Activity Pays (BAP). As we interpret it, BAP is more broadly construed to cover beneficiaries that may or may not be 'innocent' as well as agents and activities that have not produced actual harm but only a risk of harm.

Whether it is fair that beneficiaries pay in cases where risk-initiators or contributors to the problem for some reason cannot pay – for instance because they no longer exist – have been thoroughly examined in the literature on climate justice (e.g. Shue 1999, Neumayer 2000, Caney 2005, Page 2008). What to our knowledge is less discussed, is to what extent positive contributions by risk-initiators to the welfare of others can give desert-based reasons to invoke BAP, even when the risk-initiators still exist and could bear the costs of precautions. It might be seen as fair in light of desert that the risk-initiator does not bear all the costs of something that others also benefit from. That is, the positive contributions stemming from the activity that

introduces the risk might be a reason to reduce the burden on the risk initiator to fully shoulder the costs of precautions.

A fourth principle can be formulated as stating that those who benefit from taking precautions should pay for their costs. We call this principle Beneficiary of Precaution Pays (BPP). In the case of the green sea turtles discussed in Section 4, BPP would imply that the environmentalists and others who want to save the endangered turtles would benefit from taking precautions and should therefore pay. Benefits of a precaution are not limited to avoiding direct harm from the activity, as a precaution might have beneficial consequences of its own, such as spurring economic or technological innovations. As in the case of BAP, considerations of desert can support using Beneficiary of Precaution Pays (BPP). However, we suggest in the next section that desert-based arguments for BPP are most plausible when combined with some further consideration, for instance, that the precaution benefits people who are relatively well off while imposing costs on those who have the least ability to pay.

Like RIP and ATP, BAP and BPP coincide in some cases, particularly when beneficiaries of the activity and beneficiaries of the precaution are largely coextensive. But the two principles diverge when some of those who would benefit from the precaution do not also benefit from the activity, as illustrated by cases discussed in Sections 3 and 4. Previous discussions of fair distribution of costs of precautions have not explicitly distinguished BAP and BPP, and consequently have not provided an account of how they should be balanced against one another. Nor has previous literature explained how BAP and BPP should be used in conjunction with RIP or ATP. Yet these principles can easily generate conflicting recommendations, as those who benefit from the activity or the precaution need not be identical to risk initiators and may not have the greatest ability to pay.

Consideration of rights add a further level of complexity. According to a conception of rights developed by Robert Nozick (2013), rights function as 'side-constraints' on the pursuit of a just outcome. In our context, this implies that rights can constrain the use of distributional

principles. For instance, invoking RIP in the case of some activity can be constrained by the risk-initiator's right to engage in it, for example because it is necessary for survival. In cases in which initiators and beneficiaries of the risk largely overlap, such circumstances may lead to a plausible argument that Beneficiary of Precaution Pays (BPP) should be invoked. Conversely, the *lack* of a right of the risk initiator to engage in the activity might strengthen the case for insisting upon RIP. And invoking ATP to make some beneficiary of precautions *B* pay for a risk initiated by *A* may be constrained by the (prima facie) right of *B* not to be exposed to risk by *A* without consenting to it (Hansson 2003). In this context, the libertarian concept of entitlement might also be relevant. If some person *A* who has a right to perform *x* is hindered by *B* in the performance of *x*, then *A* may be entitled to some form of compensation (Nozick 2013, 57-84).

The complexity of the considerations examined in this section points to the need for a systematic framework. Developing such a framework is the task of the following section.

3. A framework for distributing precautionary costs

Our framework consists of a sequence of defaults illustrated by nested circles that can be expanded when there are strong reasons for sharing burdens (Figure 1). At the core is RIP, with BAP as the next circle, BPP after that, and in the outer circle a general responsibility of others who are not affected by the activity to shoulder the costs of precaution ('others pay'). In this framework, ATP and considerations of desert and rights function as reasons for decisions about whether or not to broaden responsibility for sharing costs.

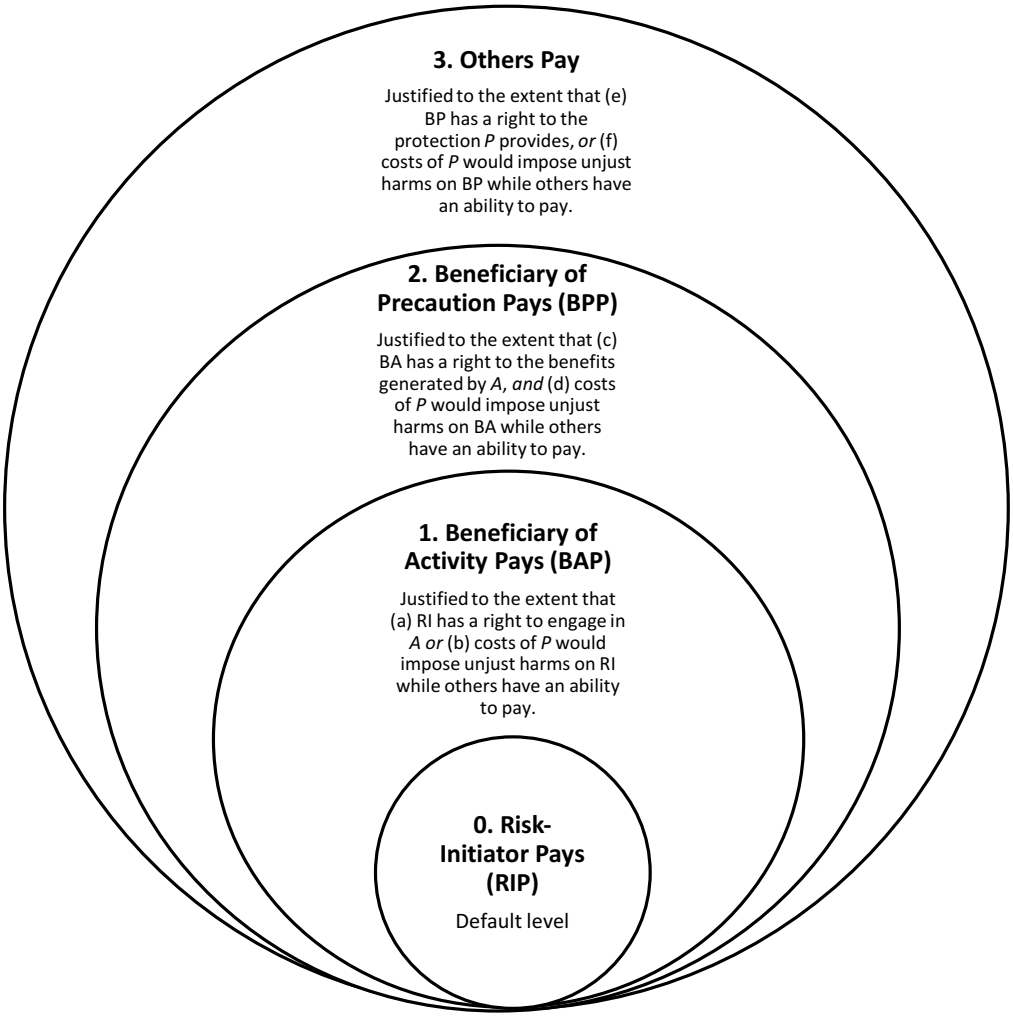

Figure 1. The figure can be used by decision makers as a heuristic framework for analysing the fairness of alternative distributions. An unjust harm involves a violation of rights or distributive or procedural justice. Abbreviations: RI=risk-initiator; BA=beneficiary of activity; BP=beneficiary of precaution; *A*=activity; *P*=precautionary measure.

Let us consider the rationale for this proposal, beginning with reasons for the role of RIP as the default starting point. There are several pragmatic reasons for this choice. Making the risk-initiator the default bearer of the cost of precautions provides a built-in incentive to avoid

activities that unnecessarily impose risks on others. When faced with the prospect of paying the cost of a precaution, the risk initiator has an incentive to ask if the risk is worth taking at all, or if the activity can be modified to mitigate that risk (e.g. through a redesign of production processes that avoids reliance on a hazardous material). In such circumstances, the risk-initiator will proceed with the activity only if there is some benefit, such as profit, to be gained from it that exceeds the costs of precautions. In addition, the risk-initiator is often in the best position to carry out precautions and to do so in a timely manner.

Treating RIP as the default can also be supported by reflections on the ethics of risk impositions. The special responsibility of the risk initiator to reduce the risk when required is related to the fact that the risk is *imposed* on someone by the risk initiator. Now, if the risk does not materialize, then no one is actually harmed by the activity. But that does not mean that no harm is done. Being at risk can itself be harmful (Nozick 2013, Hayenhjelm and Wolff 2012). Furthermore, a belief on the part of others that you are at risk may also entail substantial harms, because it may affect others' behaviour towards you in ways that negatively impact your welfare. For example, the economic value of your property may be significantly diminished if others believe it is at risk of toxic contamination from a nearby chemical factory.

Consideration of other principles of distributive justice discussed in Section 2 reinforce the role of RIP as the default, and they also help to guide decisions about how costs should be distributed as one expands the circle. Let us explore this systematically by considering two types of cases separately. In the first case, the risk initiators and the beneficiaries of the activity are coextensive, and in the second some beneficiaries of the activity are not risk initiators. To illustrate the first case, consider a person who performs chemical experiments involving explosive materials in his basement as a hobby, thereby imposing risks on his neighbours. The neighbours make no contribution to the risk, and do not benefit from it, while the risk initiator – the would-be chemist – benefits by being able to engage in an activity he enjoys.

Requiring that the cost of precautions (e.g. transforming the would-be chemist's basement into a chemical laboratory with all of the required safety apparatus) be borne by the neighbours, then, would be an instance of BPP. There are several reasons against making the beneficiary of the precaution pay in this case. Dangerous chemical experiments are not something one has a right to undertake in a private residence. Furthermore, making the neighbours pay would be unjustified from a desert perspective, since the chemistry experiments make no contribution to the wellbeing of the neighbours. Thus, while bearing the cost of precautions may be harm for the would-be chemist, there is no plausible argument that it is an *unjust* harm (as defined in Figure 1).

In cases where the beneficiary of the activity is not identical to the risk initiator, arguments that others besides the risk initiator should contribute to paying costs of precaution can be supported by considerations of just deserts. For example, consider a chemical industry that is the largest employer in a region wherein everyone benefits economically from the industry to varying degrees. In this case, there is a stronger argument from desert that others besides the risk initiator should share the costs of precautions. Given the economic contributions of the chemical industry in this example, there is a plausible argument that it would be unjust for industry to be the sole bearer of costs of precautions (i.e. an 'unjust harm' in the sense of Figure 1). Thus, funds for agencies that regulate the chemical industry could be supported by taxes from the general public as well as taxes on industry. Of course, to what extent the chemical industry deserves social support in bearing the costs of precautions depends, inter alia, on how equitably the economic benefits are distributed. If these are highly concentrated in a small capitalist class, then both desert and ATP suggest that industry should bear the bulk of the costs. To the extent that benefits are distributed more equally, the risk initiator has a stronger moral basis for claiming that other beneficiaries of the activity should also pay.

As displayed in Figure 1, we suggest that the next steps after RIP are BAP and then BPP. Why should the beneficiary of the activity have a greater obligation to pay for precautions than

the beneficiary of precautions? Take, again, the example of the chemical factories that are the largest employer in a region. As before, the chemical factories and their owners are the risk-initiators and all inhabitants of the region benefit from its economic output, either directly or indirectly. However, suppose that there are, in addition, other regions or countries that enjoy little or no gain from the chemical factories but suffer from their adverse environmental effects, for instance, in the form of pollution to air or water. In this case, some beneficiaries of the precaution are not also beneficiaries of the activity. A natural intuition here is that the default should be that the beneficiaries of the activity should pay before those who only benefit from the precaution. Why?

One plausible reason is that the risk initiator can give, subject to qualifications noted above, desert-based reasons to other beneficiaries of the activity that they should shoulder part of the burden of the precautions. But the risk initiator can give no such reason to people in other states or regions that are subject to the risks of the activity but do not enjoy its benefits. Moreover, pragmatic reasons similar to the case of RIP apply here as well. If BAP kicks in before BPP, then the beneficiaries of the activity must consider whether its social benefits are worth the costs of precautions. This may prompt them to reconsider engaging in the activity, or to explore ways in which the activity can continue but with mitigated risks. In contrast, placing BPP before BAP in the circle would encourage risk impositions that are not justified by their social benefits.

Nevertheless, in some cases there may be legitimate reasons for invoking BPP. Consider two cases: the first in which the beneficiaries of the activity and the beneficiaries of the precaution are coextensive, and the second in which they are not. In the first case, if BAP is applicable (e.g. for reasons of desert), then so too is BPP. (One might ask whether there might be others who should pay, such as the international community, but we delay this question until later.) In the second case, can there be grounds for insisting that those who benefit from the precaution but not the activity should contribute to costs of the precaution? Here rights to engage in the activity and ATP are relevant. In general, if those who benefit from the activity

have a right to those benefits (e.g. because they are necessary for survival) but are unable to bear the costs of precautions, then there may be grounds for invoking BPP. The turtle fishing example discussed in Section 4 illustrates this pattern.

The final ring in Figure 1 is 'others pay,' in which responsibility for costs of precaution is borne, at least partially, by those not at risk from the activity nor involved as initiators or beneficiaries. In our framework, 'others pay' is considered last. But why should we place BPP before those who are unaffected by the activity in the distribution circle? We suggest that this is plausible for desert-based reasons. Those carrying out the precaution can claim to the beneficiary of precaution that they are making some contribution to their welfare (i.e. by mitigating a risk). However, they cannot make similar claims to those unaffected by the action. From a related pragmatic perspective, those with a stake in enacting the precaution have an incentive to support it that is not possessed by those who are not impacted by the activity.

To illustrate grounds for invoking 'others pay,' consider a case in which the risk initiators, beneficiary of the activity, and beneficiary of the precaution all consist of the same group. Suppose that the activity generates severe health risks, but is also a necessity of life and that the people involved are not able to afford a safer alternative. A possible example here might be burning organic materials such as dung for cooking and heating inside homes. In such a case, basic rights of sustenance and health support taking precautions (e.g. the introduction of cleaner fuels), but it may be that only those unaffected by the activity are able to pay for them.

Finally, we would like to note one type of consideration that has not been explicitly mentioned so far, namely, the existence of historical and systemic wrongs, such as colonialism or racial discrimination. Such considerations can enter in our framework in several ways, for instance, by supporting the rights of certain groups to the protections provided by the precaution or to engage in certain culturally significant practices that may generate some environmental risk. Thus, a complex set of historical, social or economic considerations may be involved in

judgments about rights to engage in an activity or as reasons for why bearing the costs of a precaution would constitute an unjust harm for some but not others.

4. Cases

In the following we present two cases showing how the framework can be applied. Case 1 is based on a study by Roland Castro (2005) on green sea turtle fishing in Costa Rica. It illustrates how expanding the circle from RIP to BPP can be justified according to our framework. Case 2, about the prospect of deep sea mining in Papua New Guinea, illustrates a situation where there are strong reasons to remain at the default level of RIP.

Case 1: Green sea turtles in Costa Rica

Because of their strategic location in the Central American isthmus, Costa Rican shores host nesting populations of five of the seven existing species of sea turtle. Carribean people have traditionally hunted the green sea turtle (*Chelonia mydas*) for eggs, fat and meat. As a consequence of the hunting, the turtle population is believed to have come close to extinction in the 1960s. Some estimates indicate that nearly every female turtle arriving to nest in the area, which is now known as the Tortuguero National Park, was captured to make turtle soup for the export market, and for meat and eggs for the local market (Castro 2005).[79] To counter this activity, the Costa Rican Government enacted a regulation in 1982 officially establishing a quota of 1800 for the annual capture of green sea turtles, as well as requiring that butchering only take place in state-regulated slaughterhouses and their meat only be sold within the country. As a consequence, the permitted level of harvest was significantly reduced. However, by the late 1990s high rates of

[79] The green sea turtle is currently listed as an endangered species in the IUCN Red List of Threatened Species, available at http://www.iucnredlist.org/ (accessed 14.09.2017).

poaching meant that the number of turtles killed were many times higher than the legal limit, thereby putting unacceptable pressure on the already fragile population.

In light of this evidence, sea turtle conservation groups, environmental non-governmental organizations and some ecotourism hotels from Tortuguero requested INCOPESCA – the Costa Rican Fisheries Authority (Instituto Costarricense de Pesca y Acuacultura) – to amend this regulation and prohibit all hunting of green turtles. This attempt being unsuccessful, the groups filed a lawsuit in May of 1998 to challenge the regulation before the Constitutional Court, a branch of the Costa Rican Supreme Court.

In light of the uncertainty about how the hunting regime under INCOPESCA was affecting the ecological equilibrium of the species, the petitioners invoked the precautionary principle and asked the Court to annul the regulation to prevent the extinction of the green sea turtles, emphasizing that the species was considered endangered and threatened by extinction. INCOPESCA on their part claimed that they did all they could to prevent such extinction through implementing the regulation. They based their defence on the argument that no scientific evidence was available that could prove that the species was facing extinction under the current regime (exactly the type of argument that the precautionary principle is designed to counter).

The General Attorney's Office and the *Amicus Curioae* brief presented by the Costa Rican Ombudsman supported the petitioners' case. The Ombudsman invoked the precautionary principle against the regulation permitting turtle hunting. The Constitutional Court issued its decision on 19 February 1999, ruling in favour of annulling the regulation. Subsequently, INCOPESCA published a resolution stating that hunting and commerce of the green sea turtle were prohibited, thereby officially ending the practice.

From a socioeconomic perspective, the ban on harvesting has had its winners and losers. While the villagers in Tortuguero in the end can be said to have gained economically from the prohibition, because of the positive effects it had on tourism (Troëng, Chamorro, and Silman 2002), the fishermen who lost out were mostly from the Port of Limón (Castro 2005). Not

benefitting to any significant extent from the tourism in Tortuguero, these fishermen seemingly lost their livelihood without compensation.

Let us now look at how this process can be analysed and evaluated according to our framework. For the most part, the risk-initiators – in this case the turtle fishermen – have been compensated for the burdens (consisting mostly of opportunity costs and costs from retraining for new professions) of taking precautions in this case, through programmes and initiatives by the Caribbean Conservation Corporation (CCC) and the National Park Service. Both these organisations could be categorized as beneficiaries of precautions according to our framework. The beneficiaries of the activity (turtle fishing) are mostly coextensive with the risk-initiators, so that the question in this case is whether to expand the circle to BPP. According to our framework, there are two main reasons for doing so. The first is that taking away the fishermen's livelihood may be viewed as a violation of their basic right of subsistence (Shue 1996). Second, ATP gives us a reason to widen the circle, since presumably organisations such as CCC and the National Park Service have a greater ability to pay than local fishermen and villagers. It could further be argued that taking away the livelihood of the fishermen would risk pushing them below an acceptable standard of living – e.g. beneath the poverty line as defined by the World Bank – which implies that ATP can be invoked on sufficientarian grounds. Thus, there is a case to be made that banning sea turtle fishing without any form of compensation for the fishermen would have been an unjust harm.

Finally, it could be asked whether the former turtle fishermen of Port of Limón, who did not benefit from the thriving tourist industry in Tortuguero, should not also have received some form of compensation for the burdens they had to bear from the precautionary ban on turtle fishing. This could be argued on egalitarian grounds, in particular if it cannot be shown that these fishermen, presumably being the worse off group, benefit from the inequality between themselves and the Tortuguero fishermen (cf. the difference principle). Their right to subsistence may be said to have been violated in the same manner as the right of the Tortuguerans. If there is

no reason to discriminate other than that of the geographical location, this may be a reason for compensating the fishermen, perhaps by redistributing some of the benefits gained from tourism or helping them take part in the tourist industry.

Case 2: Deep sea mining in Papua New Guinea

Deep sea mining involves retrieving minerals such as copper, gold, silver and zinc from the ocean floor at great depths. Several authors have recommended a precautionary approach to deep sea mining, which to date has not been carried out on a commercial scale (Halfar and Fujita 2002, Wedding et al. 2015, Mengerink et al. 2014). In the following we discuss the Solwara 1 mining project proposed by Nautilus Minerals Inc. in the Bismarck Sea, off the cost of Papua New Guinea (PNG).[80]

The onshore and offshore components of Solwara 1 would be in the provinces of East New Britain and New Ireland. Rosenbaum (2011)[81] argues that communities in both provinces will face 'a range of significant risks related to the project' (Rosenbaum 2011, 22). The Bismarck Sea underpins local culture and provides food and economic livelihoods for surrounding coastal communities. Further research is needed to determine the effects of the Solwara 1 project on subsistence fishing around the Bismarck Sea and family livelihoods. However, according to Rosenbaum the environmental impacts described in Nautilus's own Environmental Impact Statement (EIS) (Nautilus Minerals 2008) indicates that Solwara 1 has the potential to 'erode the long term economic base of local communities' (Rosenbaum 2011, 22).[82] Mining activities may also exacerbate social problems already faced by island communities, as acknowledged by Nautilus in their EIS (Nautilus Minerals 2008). Moreover, it is possible that the project may

[80] The precautionary principle is recommended applied in the PNG case by Birney et al. (2006).
[81] The description in the current paper of potential socio-economic impacts of the Solwara 1 project is based mainly on Rosenbaum (2011), which was published with support from MiningWatch Canada, CELCoR (The Centre for Environmental Law and Community Rights Papua New Guinea), Oxfam Australia, and The Packard Foundation. Rosenbaum is affiliated with the Deep Sea Mining Campaign. See http://www.deepseaminingoutofourdepth.org/ (accessed 20.04.2017). For further discussion, see Sing (2015), Filer and Gabriel (2018).
[82] For a report on risks and uncertainties associated with deep sea mining, see ECORYS (2014).

affect spiritual connections between local communities and the marine environment (Rosenbaum 2011). Finally, Rosenbaum argues that the Solwara 1 mine may also impact national tuna fisheries, potentially creating health risks for people living in the villages and towns in the vicinity of the Bismarck Sea.

Nautilus and the Government of PNG argue that Solwara 1 will bring significant benefits to PNG. The Nautilus EIS states that the project will probably generate revenues in excess of US$1 billion, as well as 140 jobs. However, total tax, duties and royalty payments to the Government of PNG are estimated at only US$40.8 million over the nominal life of the project (Nautilus Minerals 2008, 10-4). Furthermore, the community development fund to be established by Nautilus to support local health and education projects represents a relatively small proportion of the revenues. Nautilus will contribute two PKG kina for every tonne of ore mined, providing approximately PGK5.8 million (about US$1.8 million) over the life time of the project (Nautilus Minerals 2008, 10-5).

The Government of PNG has reserved the right to a 15% joint venture partnership in the Solwara 1 project.[83] While such an arrangement may secure a greater revenue stream for the Government, Rosenbaum argues that it 'would represent a gross conflict of interest that would compromise the PNG Government's capacity to regulate the mining activity' (Rosenbaum 2011, 25). Moreover, she argues, experience demonstrates that 'the lack of good governance and accountability means that revenues accrued by the Government of Papua New Guinea may not necessarily translate into benefits for citizens' (Rosenbaum 2011, 25).

The socio-economic impacts described here are of crucial importance to evaluating the distributional consequences of taking precautions in the Solwara 1 case. Importantly, if it is correct that revenues accrued by the Government of Papua New Guinea are unlikely to translate into benefits for its citizens, then this is a reason to say that the Government of PNG should be separated from the general populace at the levels of RIP, BAP and BPP in our framework. It also

[83] http://www.deepseaminingoutofourdepth.org/3366/ (accessed 20.04.2017).

provides reason to believe that the ability to pay of the citizens of PNG does not reflect the ability to pay of the Government of PNG.

This suggests that there is no need in this case to go beyond the default stage of RIP. It may, as indicated, be argued that both the PNG Government and Nautilus are risk-initiators. Does this not trigger ATP on the behalf of PNG, which according to UNDP is a lower to middle income country?[84] As we have seen, ATP aims to protect those who are the worst off from having to take on burdens that would make them even worse off. The worst off in this case must be said to be the people of PNG. Since the people of PNG are not (relevantly) identical to the Government of PNG in this case, ATP does not warrant expanding the circle due to the distributional consequences for those who are worst off. Finally, there seems to be no relevant rights to be claimed from the side of the risk-initiators. Rather, citizens may have their right to subsistence threatened by the mining project.

Deciding on how the burden should be shared between the Government of PNG and Nautilus would presumably be a matter of negotiation. At first glance it does not seem entirely unreasonable that the PNG Government takes its fair share. However, if this affects the citizens of PNG, then ATP and doubts about whether the citizens of PNG will benefit – on balance – from the mining suggest that the Government should not pay for precautions.

5. Conclusion

The aim of this paper has been to develop a framework that can aid decisions about how to distribute costs of taking precautions against environmental threats. The framework can be used in situations where the precautionary principle is applied. Moreover, it can be used to address distributional issues arising from taking precautions in the more general sense referred to in the

[84] United Nations Development Program (UNDP) ranks PNG as a lower to middle income country with a gross national per capita income of US$2,386. See http://www.pg.undp.org/content/papua_new_guinea/en/home/countryinfo.html (accessed 27.03.2017).

introduction, for instance in cases where a cost-benefit approach is applied. If a cost-benefit analysis reaches the conclusion that costly measures should be taken against some risk, then – as when applying the precautionary principle – the ethical question arises of how the costs should be distributed. Our framework can be used to analyse the fairness of alternative distributions.

It should be noted that there are some general problems of distribution that are not discussed in our paper. A well-known problem is the so-called index problem, which arises from the difficulty of *measuring* the costs (and benefits) to be distributed (Lamont and Favor 2016). This problem is less pertinent in cases where what is to be distributed are economic costs, since the measurement problem arises first and foremost because of a difficulty in finding a common value measure, or in commensurating values of different qualities. Another problem arises from difficulties in defining the right time frame for the distribution, and in comparing time frames (future vs. present costs, for instance) (Lamont and Favor 2016). Moreover, the so-called non-identity problem can make it difficult to evaluate costs of precautions with regard to future generations (Parfit 1982, 1984), and the notion of historical responsibility poses a challenge with regard to risks initiated by individuals or groups in the past (e.g. Page 2008). While such problems do not preclude equitably distributing costs of precautions, they are serious theoretical problems – with potentially serious practical implications – that decision-makers should be aware of and that call for further research.

Epilogue

Should We Mine?

Modern societies need minerals. Metals found in rich ocean floor deposits are crucial for smart electronics and green technologies, such as electric cars and solar power. As the Global Sustainable Development Report stresses, green technologies are critical for sustainable development and play a key role in attempts to mitigate catastrophic impacts of climate change on nature and society (United Nations 2016).

Well-functioning societies and ecosystems are necessary for human well-being and flourishing (welfare). Human welfare is morally important – it is a basic *normative factor* in considering whether deep sea mining should be conducted.[85] However, human welfare is a complex phenomenon. Material wealth is essential to our welfare, and we value the benefits we get from nature in the form of resources and services. However, many of us also value nature for its own sake. We may prefer to live in a world where species and ecosystems are allowed to be, to exist, even if they are of no use to humans. We may also believe we have moral duties to protect and preserve nature, which do not trace back to human needs or preferences.

Sound evaluation of deep sea mining requires balancing relevant normative factors. Specifically, we have to balance preferences and needs for material benefits from mining with factors such as *existence value* and (presumably) *inherent value*. How should this be done? As a first approximation, we can ask what logical form such balancing should have. Suppose that we face a choice between mining and protecting a hydrothermal vent environment from destruction by mining. A conjecture could be that the balancing of normative factors in this choice would take something like the following form:

[85] I use italics in the current section to denote central concepts explained and discussed in the dissertation.

> An option *O* with the value *V* – where *V* is the value of a set of normative factors – can be chosen over any alternative *A* with a value less than or equal to *V*.

In essence, this principle says that we should choose the option with the greatest value – or, if our options are of equal value, any option. It is a version of a standard principle of rational choice, namely *optimizing*. Is this principle sound?

In general, we seem justified in choosing *O* in cases where *O* is clearly better than *A* – i.e. in cases of *nominal-notable* options. For instance, if (1) the relevant normative factors for deciding between mining and protecting the vent environment are societal benefits and avoidance of environmental harm, (2) *O* is a mining project involving substantial societal benefits and very little environmental harm, and (3) *A* is an ineffective and costly protection plan – then we seem justified in choosing *O*. As shown in 'Hard Environmental Choices', however, we cannot assume that if no option is considered better than any other, then our options are of equal value and we are justified in choosing any option (for instance by an arbitrary procedure such as coin-tossing). What the study suggests, is that there are cases where no option is worse, better or equal to any other with respect to the relevant choice criteria. I believe that these cases – called cases of *parity* in the dissertation – are very common, and that they explain many of our dilemmas in environmental decision-making.

Deciding whether a particular case is one of nominal-notable options or one of parity can be difficult. It involves the complicated task of specifying the values involved, or assigning the right values or weights to the relevant normative factors. Given the uncertainties of deep sea mining, we should expect that many cases will leave us in the dark. However, with continued research efforts from natural sciences, engineering, economics, philosophy, and other relevant fields, the basis for understanding the nature and normative status of particular cases of deep sea

mining will (presumably) improve, making it easier to identify and adequately describe the cases of parity, or the hard cases.

Once a case is considered a hard one, the question of *moral identity* is brought to the fore. The argument from moral identity implies that in hard cases, whether or not we should mine may be a question of what kind of people or society we want to be, or think we should be. This is a deep and fundamental philosophical question, which may have no expert answer. In other words, the question of moral identity points to the need for deliberative processes involving not only experts in relevant scientific and academic fields, but also the wider public. It points to the need for what RRI parlance refers to as 'inclusive governance' (Stilgoe, Owen, and Macnaghten 2013), for instance on the basis of public debate about whether we want to be a society that profits from environmentally harmful mining on the ocean floor, or a society that allows duties of respect for nature (or something similar) to outweigh material benefits in cases of parity.

Allowing a certain amount of deep sea mining may prove to be the right course of action in the face of mineral shortages and climate change – especially if it can replace some of the very harmful and increasingly ineffective terrestrial mining of today. But this scenario seems likely only if research and innovation on mining and alternatives to mining, such as recycling, continue in a scientifically sound and responsible manner. This may involve adopting a precautionary approach. One precautionary measure could be to continue exploration, while halting commercial extraction. Whatever precautionary measures are found reasonable from a scientific point of view, the framework developed in 'Precaution and Fairness' can help ensure that they are implemented in a fair and effective manner, with regard to the interests of industry and the general populace alike.

Further research

Suggestions for further research regarding the particular problems addressed in the studies have been given throughout the dissertation. In the current section, I make some additional

suggestions. Moreover, I draw attention to some questions regarding deep sea mining which are not taken up by the dissertation, but which could benefit from further philosophical discussion.

Inherent value

Whether and in what way it makes sense to speak about the intrinsic or inherent value of nature has been debated by environmental philosophers and other scientists and academics for decades. The question is important because, arguably, how we conceive of the intrinsic or inherent value of nature may influence the weight given to non-anthropocentric concerns in decisions and evaluations. I would suggest that in the context of deep sea mining, the issue of intrinsic or inherent value should be discussed with attention to the particular phenomenology of hydrothermal vent environments. We will probably risk losing species with which we have had no experience or encounter. Some species may disappear before we knew they existed. In general, the environments involved are far removed from our direct experience. In what way do these facts about the phenomenology of the hydrothermal vent environments affect the value we consider the species and organisms inhabiting them to have? Does it matter whether humans could *potentially* have experiences with or preferences with regard to these species and organisms? Answering such questions may provide us with a better understanding of the value and moral status of deep sea creatures and environments.

Dilemmas of conservation

Conservation biology is concerned with the viability of species and ecosystems impacted by human activities (Sandler 2012). Different strategies have been suggested by conservation biologists to protect endangered species and ecosystems, or to restore environments damaged by human activities. As Sandler (2012) points out, however, such strategies may give rise to ethical dilemmas. In particular, Sandler points out a dilemma posed by macro-scale anthropogenic

ecological change, and global climate change in particular, for *in situ* (place-based) preservation and native species prioritization. As the global climate changes, ecosystems change; some species may not survive the changes and others may thrive. Such developments, Sandler argues, bring the goal of species preservation into tension with some of the basic normative postulates of conservation biology. Under these conditions, applying *in situ* preservation and native species prioritization may require altering ecosystems in ways that may threaten species inhabiting the 'new' ecosystem.

> Moreover, the greater the novelty of a system, the less naturalness (understood as independence from human impacts), natural historical continuity, and native species there are in the system to be preserved. Thus, in many places restoration and preservation already are decreasingly viable. Or, in order to be viable, they must be diminished. Restorations must have less historical fidelity and preservation will involve fewer native species and less historical independence (Sandler 2012, 56).

Global climate change, Sandler stresses, promises to further undermine *in situ* preservation and ecological restoration, since it will 'generate even greater novelty on an even greater scale' (Sandler 2012, 56). To what extent is this a dilemma with regard to the protection of hydrothermal vents? How should conservationists respond to such dilemmas in the context of deep sea mining?

Ethics of precaution

The dissertation addresses the issue of how costs of precautions taken against environmental degradation from mining should be distributed. It does not, however, discuss in detail the rationale for taking precautionary measures. This issue deserves closer examination. For instance, it could be interesting to discuss reasons for taking precautionary measures in light of the issue of

moral identity. It is suggested in 'Hard Environmental Choices' that our choice of approach can be affected by moral identity. What kind of moral identities are at issue in discussions about the precautionary principle? How can we evaluate their significance for decision-making? These issues are discussed briefly in 'Hard Environmental Choices', but would be interesting to explore in more detail.

Common heritage of mankind

An important topic that is not discussed in the dissertation is how the (potential) *benefits* of deep sea mining can be fairly distributed. This is a particularly interesting question when considering mining in what is called 'the Area', which is the part of the seafloor not included in any country's national jurisdiction or exclusive economic zone (EEZ). United Nations Law of the Sea Treaty (UNCLOS) declares the Area and its resources to be the 'common heritage of mankind'.[86] What exactly are the normative implications of this idea for the distribution of economic and other gains from deep sea mining? How does the idea relate to different conceptions of justice in the philosophical literature? Does the principle of the common heritage of mankind have an adequate rationale, or should it be replaced by a different principle?

The need for minerals

An argument for many forms of mining, including on the ocean floor, is that society needs minerals (see for instance the beginning of this chapter). This argument should be scrutinized philosophically. Importantly, we should examine the nature of this 'need'. To what extent are basic needs at stake, in contrast to less pressing needs and preferences, such as preferences regarding comfortable living or luxury goods? Philosophical analysis could help clarify how pressing the need for minerals in modern societies is.

[86] UNCLOS, Part XI, Section 2, Article 136. See: http://www.un.org/Depts/los/convention_agreements/texts/unclos/part11-2.htm (accessed 21.04.2019).

References

Alcamo, Joseph, Elena Bennett, Neville J. Ash, Colin D. Butler, J Baird Callicott, Doris Capistrano, and Stephen Carpenter et al. 2003. *Millennium Ecosystem Assessment: Ecosystems and Human Well-being: A Framework for Assessment*. Washington, DC: Island Press.

Anderson, Elizabeth. 1997. "Practical reason and incommensurable goods." In *Incommensurability, Incomparability, and Practical Reason*, edited by Ruth Chang, 90-109. Cambridge, Massachusetts: Harvard University Press.

Aristotle. 1976 [c. 400 BC]. *The Nicomachean Ethics*. Translated by J.A.K. Thomson. London: Penguin.

Arneson, Richard J. 2000. "Luck egalitarianism and prioritarianism." *Ethics* 110 (2):339-349.

Barry, Christian, and Robert Kirby. 2017. "Scepticism about Beneficiary Pays: A Critique." *Journal of Applied Philosophy* 34 (3):285-300. doi: 10.1111/japp.12160.

Barton, D., T. Bongard, H. Lindhjem, G. Rusch, J. Thomassen, and S Öberg. 2011. Ecosystem services – from concept to practice? A summary of presentations at the ØKOSIP seminar, January 10th, NINA Trondheim – NINA Report 673.

Bentham, Jeremy. 1789. *Introduction to the Principles of Morals and Legislation*. London: Printed for T. Payne, and Son, at the Mews Gate.

Berlin, Isiah. 1990. *The Crooked Timber of Humanity: Chapters in the History of Ideas*. London: John Murray.

Berlin, Isiah. 2002. *Liberty*. Edited by Henry Hardy. Oxford and New York: Oxford University Press.

Birney, Kristi, Amber Griffin, Jonathan Gwiazda, Johnny Kefauver, Takehiko Nagai, and Douglas Varchol. 2006. "Potential deep-sea mining of seafloor massive sulfides: A case study in Papua New Guinea." Donald Bren School of Science and Management.

Boot, Martijn. 2009. "Parity, incomparability and rationally justified choice." *Philosophical Studies* 146 (1):75-92.

Boschen, R. E., A. A. Rowden, M. R. Clark, and J. P. A. Gardner. 2013. "Mining of deep-sea seafloor massive sulfides: A review of the deposits, their benthic communities, impacts from mining, regulatory frameworks and management strategies." *Ocean & Coastal Management* 84:54-67. doi: http://doi.org/10.1016/j.ocecoaman.2013.07.005.

Bradley, Ben. 2007. "A paradox for some theories of welfare." *Philosophical Studies* 133 (1):45-53.

Brink, David Owen. 1989. *Moral Realism and the Foundations of Ethics*. Cambridge: Cambridge University Press.

Broome, John. 1997. "Is incommensurability vagueness?" In *Incommensurablity, Incomparability, and Practical Reason*, edited by Ruth Chang, 67-89. Cambridge, Massachusetts: Harvard University Press.

Broome, John. 1998. "Extended Preferences." In *Preferences*, edited by Ulla Wessels and Christoph Fehige, 271-287. Berlin: De Gruyter.

Brülde, Bengt. 1998. *The Human Good*. Göteborg: Acta Universitas Gothoburgensis.

Callicott, J. Baird. 1989. *In Defense of the Land Ethic: Essays in Environmental Philosophy*. Albany, N.Y: State University of New York Press.

Caney, Simon. 2005. "Cosmopolitan justice, responsibility, and global climate change." *Leiden journal of international law* 18 (4):747-775.

Castro, Roland. 2005. "Protection of Sea Turtles: Putting the Precautionary Principle into Practice." In *Biodiversity and the Precautionary Principle: Risk and Uncertainty in Conservation and Sustainable Use*, edited by Rosie Cooney and Barney Dickson. London: Earthscan.

Ceballos, Gerardo, Paul R Ehrlich, Anthony D Barnosky, Andrés García, Robert M Pringle, and Todd M Palmer. 2015. "Accelerated modern human–induced species losses: Entering the sixth mass extinction." *Science advances* 1 (5):e1400253.

Chang, Ruth. 1997. "Introduction." In *Incommensurability, Incomparability, and Practical Reason*, edited by Ruth Chang, 1-34. Cambridge, Massachusetts: Harvard University Press.

Chang, Ruth. 2002. "The possibility of parity." *Ethics* 112 (4):659-688.

Chang, Ruth. 2005. "Parity, interval value, and choice." *Ethics* 115 (2):331-350.

Chang, Ruth. 2009a. "Reflections on the Reasonable and the Rational in Conflict Resolution." Aristotelian Society Supplementary Volume LXXXIII.

Chang, Ruth. 2009b. "Voluntarist reasons and the sources of normativity." In *Reasons for Action*, edited by David Sobel and Steven Wall, 243-71. Cambridge: Cambridge University Press.

Chang, Ruth. 2012. "Are hard choices cases of incomparability?" *Philosophical issues* 22 (1):106-126.

Chang, Ruth. 2013. "Grounding practical normativity: going hybrid." *Philosophical Studies* 164 (1):163-187.

Chang, Ruth. 2014. *Making Comparisons Count*. New York: Routledge.

Cohen, Gerald A. 1996. "Reason, Humanity, and the Moral Law." In *The Sources of Normativity*, edited by Onora O'Neill, 167-188. Cambridge: Cambridge University Press.

Collingridge, David. 1980. *The Social Control of Technology*. London: Frances Printer Ltd.

Crane, J. and R. L. Sandler. 2011. "Species Concepts and Natural Goodness". In *Carving Nature at its Joints: Themes in Contemporary Philosophy*, vol. 8, edited by J. K. Campbell, M O'Rourke, and M. Slater, 289-311. Cambridge, MA: MIT Press.

Donagan, Alan. 1977. *The Theory of Morality*. Chicago: University of Chicago Press.

Durden, Jennifer M, Kevin Murphy, Aline Jaeckel, Cindy L Van Dover, Sabine Christiansen, Kristina Gjerde, Aleyda Ortega, and Daniel OB Jones. 2017. "A procedural framework for robust environmental management of deep-sea mining projects using a conceptual model." *Marine Policy* 84:193-201.

Earth Economics. 2015. Environmental and Social Benchmarking Analysis of Nautilus Minerals Inc. Solwara 1 project.

ECORYS. 2014. Study to investigate state of knowledge of deep sea mining. Final report Annex 6 Environmental Analysis. FWC MARE/2012/06 – SC E1/2013/04.

Elster, Jon. 1983. *Sour Grapes*. Cambridge: Cambridge University Press.

Filer, Colin, and Jennifer Gabriel. 2018. "How could Nautilus Minerals get a social licence to operate the world's first deep sea mine?" *Marine Policy*. doi: http://dx.doi.org/10.1016/j.marpol.2016.12.001 (accessed 30.03.2017).

Frankfurt, Harry. 1987. "Equality as a moral ideal." *Ethics* 98 (1):21-43.

Gaines, Sanford E. 1991. "The polluter-pays principle: from economic equity to environmental ethos." *Tex. Int'l LJ* 26:463.

Gollner, Sabine, Stefanie Kaiser, Lena Menzel, Daniel OB Jones, Alastair Brown, Nelia C Mestre, Dick Van Oevelen, Lenaick Menot, Ana Colaço, and Miquel Canals. 2017. "Resilience of benthic deep-sea fauna to mining activities." *Marine environmental research* 129:76-101.

Gollner, Sabine, Maria Miljutina, and Monika Bright. 2013. "Nematode succession at deep-sea hydrothermal vents after a recent volcanic eruption with the description of two dominant species." *Organisms Diversity & Evolution* 13 (3):349-371.

Goodin, Robert E. 2013. "Disgorging the fruits of historical wrongdoing." *American Political Science Review*:478-491.

Goodin, Robert E, and Christian Barry. 2014. "Benefiting from the Wrongdoing of Others." *Journal of Applied Philosophy* 31 (4):363-376.

Greco, Salvatore, J Figueira, and M Ehrgott. 2005. *Multiple Criteria Decision Analysis: State of the Art Surveys*. 2 ed. New York: Springer.

Griffin, James. 1986. *Well-Being: Its Meaning, Measurement, and Moral Importance*. Oxford: Clarendon Press.

Griffin, James. 1997. "Incommensurability: What's the problem." In *Incommensurability, Incomparability, and Practical Reason*, edited by Ruth Chang, 35-51. Cambridge, Massachusetts: Harvard University Press.

Grigg, Richard W, A Malaboff, EH Chave, and J Landahl. 1987. "Seamount benthic ecology and potential environmental impact from manganese crust mining in Hawaii." *Seamounts, islands, and atolls* 43:379-390.

Halfar, J., and R.M. Fujita. 2002. "Precautionary management of deep sea mining." *Marine Policy* (26):103–106.

Hansson, Sven Ove. 2018. Risk. In *The Stanford Encyclopedia of Philosophy*. Fall 2018 Edition, URL=<https://plato.stanford.edu/archives/fall2018/entries/risk/>

Hargrove, Eugene C. 1992. "Weak anthropocentric intrinsic value." *The Monist* 75 (2):183-207.

Hausman, Daniel M, and Michael S McPherson. 2009. "Preference satisfaction and welfare economics." *Economics & Philosophy* 25 (1):1-25.

Hauton, Chris, Alastair Brown, Sven Thatje, Nélia C Mestre, Maria J Bebianno, Inês Martins, Raul Bettencourt, Miquel Canals, Anna Sanchez-Vidal, and Bruce Shillito. 2017. "Identifying Toxic Impacts of Metals Potentially Released during Deep-Sea Mining—A Synthesis of the Challenges to Quantifying Risk." *Frontiers in Marine Science* 4:368.

Hayenhjelm, Madeleine, and Jonathan Wolff. 2012. "The moral problem of risk impositions: A survey of the literature." *European Journal of Philosophy* 20 (S1):E26-E51.

Heathwood, Chris. 2005. "The problem of defective desires." *Australasian Journal of Philosophy* 83 (4):487-504.

Hsieh, Nien-hê. 2007. "Is incomparability a problem for anyone?" *Economics & Philosophy* 23 (1):65-80.

Irvine, Katherine N, Liz O'Brien, Neil Ravenscroft, Nigel Cooper, Mark Everard, Ioan Fazey, Mark S Reed, and Jasper O Kenter. 2016. "Ecosystem services and the idea of shared values." *Ecosystem Services* 21:184-193.

Jamieson, Dale. 2008. *Ethics and the Environment: An Introduction*. Cambridge: Cambridge University Press.

Jonas, Hans. 1966. *The Phenomenon of Life: Toward a Philosophy of Biology*. Chicago: Northwestern University Press.

Jonas, Hans. 1979. *Das Prizip Verantwortung*. Frankfurt am Main: Suhrkamp.

Justus, James, Mark Colyvan, Helen Regan, and Lynn Maguire. 2009. "Buying into conservation: intrinsic versus instrumental value." *Trends in Ecology & Evolution* 24 (4):187-191.

Kagan, Shelly. 1998. *Normative Ethics*. Boulder, Colorado: Westview Press.

Kant, Immanuel. 1974 [1785/1786]. *Grundlegung der Metaphysik der Sitten*. Edited by Wilhelm Weischedel. Vol. VII, *Immanuel Kant Werkausgabe*. Frankfurt am Main: Suhrkamp.

Kant, Immanuel. 1974 [1788]. *Kritik der Praktischen Vernunft*. Edited by Wilhelm Weischedel. Vol. VII, *Immanuel Kant Werkausgabe*. Frankfurt am Main: Suhrkamp.

Kenter, Jasper O, Rosalind Bryce, Michael Christie, Nigel Cooper, Neal Hockley, Katherine N Irvine, Ioan Fazey, Liz O'Brien, Johanne Orchard-Webb, and Neil Ravenscroft. 2016. "Shared values and deliberative valuation: Future directions." *Ecosystem services* 21:358-371.

Kenter, Jasper O, Liz O'Brien, Neal Hockley, Neil Ravenscroft, Ioan Fazey, Katherine N Irvine, Mark S Reed, Michael Christie, Emily Brady, and Rosalind Bryce. 2015. "What are shared and social values of ecosystems?" *Ecological Economics* 111:86-99.

Korsgaard, Christine M. 1996. *The Sources of Normativity*. Cambridge: Cambridge University Press.

Kramer, Koen, Hans L Zaaijer, and Marcel F Verweij. 2017. "The precautionary principle and the tolerability of blood transfusion risks." *The American Journal of Bioethics* 17 (3):32-43.

Kraut, Richard. 2013. "Desire and the human good." *The American Philosophical Association Centennial Series*:255-270.

Köksalan, Murat, Jyrki Wallenius, and Stanley Zionts. 2013. "An early history of multiple criteria decision making." *Journal of Multi-Criteria Decision Analysis* 20 (1-2):87-94.

Lamont, Julian, and Christi Favor. 2016. Distributive Justice. In *The Stanford Encyclopedia of Philosophy*. Winter 2016 Edition, URL = <https://plato.stanford.edu/archives/win2016/entries/justice-distributive/>

Lawford-Smith, Holly. 2014. "Benefiting from failures to address climate change." *Journal of Applied Philosophy* 31 (4):392-404.

Leopold, Aldo. 1949. *A Sand County Almanac*. New York: Oxford University Press.

Levin, Lisa A, Guillermo F Mendoza, Talina Konotchick, and Raymond Lee. 2009. "Macrobenthos community structure and trophic relationships within active and inactive Pacific hydrothermal sediments." *Deep Sea Research Part II: Topical Studies in Oceanography* 56 (19-20):1632-1648.

Light, Andrew, and Holmes Rolston, eds. 2002. *Environmental Ethics: An Anthology*. Malden: Blackwell Publishing.

Luce, R. Duncan, and Howard Raiffa. 1957. *Games and Decisions: Introduction and Critical Survey*. New York: Wiley.

Lucivero, Federica, Tsjalling Swierstra, and Marianne Boenink. 2011. "Assessing expectations: Towards a toolbox for an ethics of emerging technologies." *NanoEthics* 5 (2):129.

Martinez-Alier, Joan, Giuseppe Munda, and John O'Neill. 1998. "Weak comparability of values as a foundation for ecological economics." *Ecological economics* 26 (3):277-286.

Mengerink, Kathryn J, Cindy L Van Dover, Jeff Ardron, Maria Baker, Elva Escobar-Briones, Kristina Gjerde, J Anthony Koslow, Eva Ramirez-Llodra, Ana Lara-Lopez, and Dale Squires. 2014. "A call for deep-ocean stewardship." *Science* 344 (6185):696-698.

Mill, John Stuart. 2001 [1871]. "Utilitarianism." In *Classics of Moral and Political Theory*, 935-968. Indianapolis/Cambridge: Hackett Publishing Company.

Miller, David. 1976. *Social justice*. Oxford: Oxford University Press.

Miller, David. 2009. "Global justice and climate change: how should responsibilities be distributed? Parts I and II." *Tanner Lectures on Human Values* 28.

Milne, Heather. 1986. "Desert, effort and equality." *Journal of Applied Philosophy* 3 (2):235-243.

Moalic, Yann, Daniel Desbruyères, Carlos M Duarte, Alejandro F Rozenfeld, Charleyne Bachraty, and Sophie Arnaud-Haond. 2011. "Biogeography revisited with network theory: retracing the history of hydrothermal vent communities." *Systematic Biology* 61 (1):127.

Morato, Telmo, Simon D Hoyle, Valerie Allain, and Simon J Nicol. 2010. "Seamounts are hotspots of pelagic biodiversity in the open ocean." *Proceedings of the National Academy of Sciences*:200910290.

Munthe, Christian. 2011. *The price of precaution and the ethics of risk*, *The International Library of Ethics, Law and Technology*. Dordrecht: Springer Science & Business Media.

Nagel, Thomas. 1979. "The fragmentation of value." In *Mortal Questions*, 128–141. Cambridge, UK: Cambridge University Press.Nautilus Minerals. 2008. Environmental Impact Statement.

Nautilius Minerals Niugini Limited. Solwara 1 project. http://www.nautilusminerals.com/irm/content/pdf/environment-reports/Environmental%20Impact%20Statement%20-%20Main%20Report.pdf

Neumayer, Eric. 2000. "In defence of historical accountability for greenhouse gas emissions." *Ecological economics* 33 (2):185-192.

Nozick, Robert. 2013. *Anarchy, state, and utopia*. New York: Basic books.

Nussbaum, Martha, and Amartya Sen. 1993. *The Quality of Life*. Oxford: Oxford University Press.

Næss, Arne. 1990. *Ecology, Community and Lifestyle: Outline of an Ecosophy*. Cambridge: Cambridge University Press.

O'Neill, J., and C. L. Spash. 2000. "Conceptions of Value in Environmental Decision-Making." *Environmental Values* 9 (4):521-536.

O'Neill, John. 1992. "The Varieties of Instrinsic Value." *The Monist* (75 (2)):119–137.

O'Neill, John. 1993. *Ecology, Policy and Politics: Human Well-Being and the Natural World*. London: Routledge.

O'Riordan, Timothy. 1994. *Interpreting the Precautionary Principle*. London: Earthscan.

O'Hara, Timothy D, and Derek P Tittensor. 2010. "Environmental drivers of ophiuroid species richness on seamounts." *Marine Ecology* 31:26-38.

OECD. 1972. Guiding Principles Concerning International Economic Aspects of Environmental Policies, Recommendation C(72)128, adopted May 26, 1972, reprinted in 11 I.L.M. 1172 (1972).

Page, Edward A. 2008. "Distributing the burdens of climate change." *Environmental Politics* 17 (4):556-575.

Parfit, Derek. 1982. "Future generations: Further problems." *Philosophy & Public Affairs*:113-172.

Parfit, Derek. 1984. *Reasons and Persons*. Oxford: Oxford University Press.

Parfit, Derek. 2011. *On What Matters: Volume One*. Oxford: Oxford University Press.

Parks, Sarah, and John Gowdy. 2013. "What have economists learned about valuing nature? A review essay." *Ecosystem Services* 3:e1-e10.

Peterson, Martin. 2007. "On multi-attribute risk analysis." In *Risk: Philosophical Perspectives*, edited by Tim Lewens, 68–83. London: Routledge.

Peterson, Martin. 2009. *An Introduction to Decision Theory*. Cambridge: Cambridge University Press.

Pettit, Philip. 1997. "The Consequentialist Perspective." In *Three Methods of Ethics: A Debate*, edited by Marcia W. Baron, Philip Pettit and Michael Slote. Oxford: Blackwell Publishing.

Pimm, S. L., C. N. Jenkins, R. Abell, T. M. Brooks, J. L. Gittleman, L. N. Joppa, P. H. Raven, C. M. Roberts, and J. O. Sexton. 2014. "The biodiversity of species and their rates of extinction, distribution, and protection." *Science* 344 (6187). doi: 10.1126/science.1246752.

Ramirez-Llodra, Eva, Paul A Tyler, Maria C Baker, Odd Aksel Bergstad, Malcolm R Clark, Elva Escobar, Lisa A Levin, Lenaick Menot, Ashley A Rowden, and Craig R Smith. 2011. "Man and the last great wilderness: human impact on the deep sea." *PLoS one* 6 (8):e22588.

Rawls, John. 1971. *A Theory of Justice*. Cambridge, Massachusets: Harvard University Press.

Rawls, John. 1985. "Justice as fairness: Political not metaphysical." *Philosophy & Public Affairs*:223-251.

Raz, Joseph. 1986. *The Morality of Freedom*. Oxford: Clarendon Press.

Raz, Joseph. 1997. "Incommensurability and agency." In *Incommensurability, incomparability, and practical reason*, edited by Ruth Chang, 110-128. Cambridge, Massachusetts: Harvard University Press.

Regan, Donald. 1997. "Value, comparability, and choice." In *Incommensurability, Incomparability, and Practical Reason*, edited by Ruth Chang, 129-150. Cambridge, Massachusetts: Harvard University Press.

Riley, Jonathan. 1989. "Justice under capitalism." *Nomos* 31:122-162.

Rolston, Holmes. 1989. *Environmental Ethics. Values in and Duties to the Natural World*. Philadelphia: Temple University Press.

Rosenbaum, Helen. 2011. Out of our Depth: Mining the Ocean Floor in Papua New Guinea. edited by Natalie Lowrey, Christina Hill and Catherine Coumans: MiningWatch Canada, CELCoR, Oxfam Australia, The Packard Foundation.

Ross, William David. 1930. *The Right and the Good*. London: Oxford University Press.

Sadurski, Wojciech. 1985. *Giving desert its due: Social justice and legal theory*. Vol. 2: Springer Science & Business Media.

Sagoff, Mark. 1986. "Values and preferences." *Ethics* 96 (2):301-316.

Sagoff, Mark. 1988. "Some problems with environmental economics." *Environmental Ethics* 10 (1):55-74.

Sagoff, Mark. 2000. "Environmental economics and the conflation of value and benefit." *Environmental science & technology* 34 (8):1426-1432.

Sandin, Per, Martin Peterson, Sven Ove Hansson, Christina Rudén, and André Juthe. 2002. "Five charges against the precautionary principle." *Journal of Risk Research* 5 (4):287-299.

Sandler, Ronald L. 2007. *Character and Environment: A Virtue-Oriented Approach to*

Environmental Ethics. New York: Columbia University Press.

Sandler, Ronald L. 2012. *The Ethics of Species: An Introduction*. Cambridge: Cambridge University Press.

Sarch, Alexander F. 2015. "Hausman and McPherson on welfare economics and preference satisfaction theories of welfare: a critical note." *Economics & Philosophy* 31 (1):141-159.

Scanlon, Thomas. 1998. *What We Owe to Each Other*. Cambridge, Massachusets: Harvard University Press.

Schick, Frederic. 1982. "Under which descriptions." In *Utilitarianism and Beyond*, edited by Amartya Sen and Bernhard Williams. Cambridge: Cambridge University Press.

Schlacher, Thomas A, Amy R Baco, Ashley A Rowden, Timothy D O'Hara, Malcolm R Clark, Chris Kelley, and John F Dower. 2014. "Seamount benthos in a cobalt-rich crust region of the central Pacific: conservation challenges for future seabed mining." *Diversity and distributions* 20 (5):491-502.

Sen, Amartya. 1992. *Inequality Reexamined*. Oxford: Clarendon Press.

Sen, Amartya. 2011. *The Idea of Justice*. Cambridge, Massachusetts: Harvard University Press.

Sen, Amartya. 2017. *Collective Choice and Social Welfare: Expanded Edition*. UK: Penguin Books.

Shue, Henry. 1996. *Basic Rights: Subsistence, Affluence, and US Foreign Policy*. Princeton: Princeton University Press.

Shue, Henry. 1999. "Global environment and international inequality." *International affairs* 75 (3):531-545.

Sidgwick, Henry. 1981 [1907]. *The Methods of Ethics*. Seventh ed. Indianapolis, Indiana: Hackett Publishing.

Sing, Jason. 2015. "Regulating mining resource investments towards sustainable development: The case of Papua New Guinea." *The Extractive Industries and Society* 2 (1):124-131. doi: http://dx.doi.org/10.1016/j.exis.2014.11.003.

Singer, Peter. 2011. *Practical Ethics*. Third ed. Cambridge: Cambridge University Press.

Stabell, Espen D. 2017. "Constraints on the Precautionary Principle and the Problem of Uncertainty." *The American Journal of Bioethics* 17 (3):56-57.

Steel, Daniel. 2015. *Philosophy and the Precautionary Principle: Science, Evidence, and Environmental Policy*. Cambridge: Cambridge University Press.

Stilgoe, Jack, Richard Owen, and Phil Macnaghten. 2013. "Developing a framework for responsible innovation." *Research Policy* 42 (9):1568-1580.

Stocker, Michael. 1990. *Plural and Conflicting Values*. Oxford: Clarendon Press.

Sukhdev, Pavan, H Wittmer, C Nesshöver, and B Simmons. 2010. "TEEB Synthesis Report Mainstreaming the Economics of Nature." *Geneva: UNEP*.

Sumner, Leonard Wayne. 1996. *Welfare, Happiness, and Ethics*. Oxford: Clarendon Press.

Sunstein, Cass R. 2005a. "Cost-benefit analysis and the environment." *Ethics* 115 (2):351-385.

Sunstein, Cass R. 2005b. *Laws of fear: Beyond the precautionary principle*. Cambridge: Cambridge University Press.

Sylvan (Routley), Richard. 1973. "Is There a Need for a New, an Environmental, Ethic?" Proceedings of the XVth World Congress of Philosophy, Sofia, Bulgaria.

Taylor, Charles. 1989. *Sources of the Self: The Making of the Modern Identity*. Cambridge, Massachusetts: Harvard University Press.

Taylor, Paul W. 2011. *Respect for Nature: A Theory of Environmental Ethics*. Princeton: Princeton University Press.

TEEB, UNEP. 2012. Why Value the Oceans – A Discussion Paper. edited by Yannick Beaudoin and Linwood Pendleton.

Troëng, S, E Chamorro, and R Silman. 2002. "Ban and benefits: Tortuguero at 2000." Proceedings of the Twentieth Annual Symposium on Sea Turtle Biology and Conservation. US Dep. Commer. NOAA Tech. Memo. NMFS-SEFSC-477.

Trouwborst, Arie. 2006. *Precautionary Rights and Duties of States*, *Nova et Vetera Iuris Gentium*. Leiden: Martinus Nijhoff Publishers.

Tunnicliffe, Verena. 1992. "The nature and origin of the modern hydrothermal vent fauna." *Palaios*:338-350.

UNEP. 2007. Deep-sea biodiversity and ecosystems: A scoping report on their socio-economy, management and governance. UNEP World Conservation Monitoring Centre.

United Nations. 2016. Global Sustainable Development Report 2016. New York.

United Nations. 1992. "Rio Declaration on Environment and Development." Rio de Janeiro.

Valentini, Laura. 2012. "Ideal vs. Non-ideal Theory: A Conceptual Map." *Philosophy Compass* 7 (9):654-664.

Van Dover, C. L., J. A. Ardron, E. Escobar, M. Gianni, K. M. Gjerde, A. Jaeckel, D. O. B. Jones, L. A. Levin, H. J. Niner, L. Pendleton, C. R. Smith, T. Thiele, P. J. Turner, L. Watling, and P. P. E. Weaver. 2017. "Biodiversity loss from deep-sea mining." *Nature Geoscience* (10):464–465. doi: 10.1038/ngeo2983.

Van Dover, Cindy Lee. 2010. "Mining seafloor massive sulphides and biodiversity: what is at risk?" *ICES Journal of Marine Science* 68 (2):341-348.

Van Dover, Cindy Lee. 2014. "Impacts of anthropogenic disturbances at deep-sea hydrothermal vent ecosystems: a review." *Marine Environmental Research* 102:59-72.

Van Dover, Cindy Lee, Sophie Arnaud-Haond, M Gianni, S Helmreich, JA Huber, AL Jaeckel, A Metaxas, LH Pendleton, Sven Petersen, and E Ramirez-Llodra. 2018. "Scientific rationale and international obligations for protection of active hydrothermal vent ecosystems from deep-sea mining." *Marine Policy* 90:20-28.

Van Dover, Cindy Lee, and Robert R Hessler. 1990. "Spatial variation in faunal composition of hydrothermal vent communities on the East Pacific Rise and Galapagos spreading center." In *Gorda Ridge*, 253-264. Springer.

Von Wright, Georg Henrik. 1972. "The logic of preference reconsidered." *Theory and Decision* 3 (2):140.

Wedding, LM, SM Reiter, CR Smith, KM Gjerde, JN Kittinger, AM Friedlander, SD Gaines, MR Clark, AM Thurnherr, and SM Hardy. 2015. "Managing mining of the deep seabed." *Science* 349 (6244):144-145.

Williams, Bernard. 1981. "Conflicts of values." In *Moral Luck: Philosophical papers 1973-1980*. Cambridge: Cambridge University Press.

Williams, Bernard. 2006. *Ethics and the Limits of Philosophy*. London: Routledge.

World Commission on Environment and Development. 1987. *Our Common Future*. Oxford: Oxford University Press.

Zhang, Likui, Manyu Kang, Jiajun Xu, Jian Xu, Yinjie Shuai, Xiaojian Zhou, Zhihui Yang, and Kesen Ma. 2016. "Bacterial and archaeal communities in the deep-sea sediments of inactive hydrothermal vents in the Southwest India Ridge." *Scientific Reports* 6. doi: 10.1038/srep25982.

Addendum

Co-author information

Daniel Steel

School of Population and Public Health, Maurice Young Centre for Applied Ethics, University of British Columbia, Vancouver, Canada.

Co-author on Study III: 'Precaution and Fairness: A Framework for Distributing Costs of Protection from Environmental Risks'.

Publications

Stabell, Espen Dyrnes, and Daniel Steel. 2018. 'Precaution and Fairness: A Framework for Distributing Costs of Protection from Environmental Risks.' *Journal of Agricultural and Environmental Ethics* 31 (1):55-71.

The dissertaion includes a postprint of the published article (with minor corrections).