Knowledge, beliefs and regulatory preferences in environmental health controversies: the case of endocrine disrupting substances

Sander Carl Stijn Clahsen

Sander Clahsen
Knowledge, beliefs and regulatory preferences in environmental health controversies: the case of endocrine disrupting substances
ISBN: 978-90-393-7408-5

Cover design: Yolanda Koevoets
Layout: Sander Clahsen and Yolanda Koevoets
Printed by: Gildeprint Drukkerijen

The research presented in this thesis was funded by the Dutch National Institute for Public Health and the Environment (RIVM) as part of the RIVM Strategic Programme (SPR; S/124002; TARGETS project), in which expertise and innovative projects prepare RIVM to respond to future issues in health and sustainability.

Knowledge, beliefs and regulatory preferences in environmental health controversies: the case of endocrine disrupting substances

Kennis, overtuigingen en voorkeuren voor regelgeving in milieugezondheidscontroversen:
de hormoonverstorende stoffen casus
(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht
op gezag van de rector magnificus, prof. dr. H.R.B.M. Kummeling,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op

dinsdag 16 november 2021 des middags te 12.15 uur

door

Sander Carl Stijn Clahsen

geboren op 28 september 1988
te Enschede

Promotoren:
Prof. dr. A.H. Piersma
Prof. dr. E. Lebret

Copromotoren:
Dr. I. van Kamp
Dr. T.G. Vermeire

Table of contents

List of abbreviations

4-MBC	4-methylbenzylidene camphor
ACF	Advocacy Coalition Framework
ADAS	UK agricultural and environmental consultancy
AGES	Austrian Agency for Health and Food Safety
ANSES	French Agency for Food, Environmental and Occupational Health & Safety
ARERMP	Advisory roles of experts in risk management processes
BCPC	British Crop Production Council
BEUC	European Consumer Organization
BfR	German Federal Institute for Risk Assessment
BPA	Bisphenol A
BPR	European Biocidal Products Regulation
CDC	Centre for Disease Control and Prevention
CEFIC	European Chemical Industry Council
CEO	Corporate Europe Observatory
CLP	European Classification, Labelling and Packaging Regulation
CTR	Cultural Theory of Risk
DDT	Dichlorodiphenyltrichloroethane
DEHP	Bis(2-ethylhexyl) phthalate
DES	Diethylstilbestrol
EAS	Endocrine active substance
EC	European Commission
ECPA	European Crop Protection Association
EDC(s)	Endocrine disrupting chemical(s)
EDS(s)	Endocrine disrupting substance(s)
EEC	European Economic Community
EFSA	European Food Safety Authority
EMF(s)	Electromagnetic field(s)
EP	European Parliament
EPA	Environmental Protection Agency
FAO	Food and Agriculture Organization of the United Nations
FSA	Norwegian Food Safety Authority
GM	Genetic Modification/Genetically Modified
HC	Hazard Characterization
HEAL	European Health and Environment Alliance
HI	Hazard Identification
HMNC	Hofstede's Model of National Cultures
IARC	International Agency for Research on Cancer
INERIS	French National Institute for Industrial Environment and Risks
IPCS	International Program on Chemical Safety
IRGC	International Risk Governance Center

JRC	European Joint Research Centre
KEMI	Swedish Chemicals Agency
MAC	Maximum allowable concentration
MCL	Maximum contaminant level
NGO	Non-governmental organization
NICS	Hungarian National Institute of Chemical Safety
NRC	U.S. National Research Council
PAN Europe	Pesticide Action Network Europe
PARAM	(Hybrid) participatory approaches to risk assessment and management
PDAT	Pragma-Dialectical Argumentation Theory
PM	Particulate Matter
PP	Psychometric Paradigm
PPPR	European Plant Protection Products Regulation
RA	Risk Assessment
RAP	Risk Assessment Paradigm
REACH	European Registration, Evaluation, Authorization and Restriction of Chemicals Regulation
RES	French Environmental Health Network
RF EMFs	Radio Frequency Electromagnetic Fields
RIVM	Dutch National Institute for Public Health and the Environment
SARF	Social Amplification of Risk Framework
SCCS	Scientific Committee on Consumer Safety
SCENIHR	Scientific Committee on Emerging and Newly Identified Health Risks
SCHER	Scientific Committee on Health and Environmental Risks
TUKES	Finnish Safety and Chemical Agency
UBA	German or Austrian Environment Agency
UIC	French Chemical Industry Association
UNEP	United Nations Environment Program
VFA	Danish Veterinary and Food Administration
WECF	Women Engage for a Common Future
WHO	World Health Organization
WRR	The Netherlands Scientific Council for Government Policy

General introduction

SETTING THE SCENE *Science and scientists were often considered trustworthy and credible by default. However, a slow but steady shift to a sort of 'wariness by default' can be observed. Movements denouncing human vaccination programs ('antivaxxers') have gained global momentum and the roll out of 5G radiofrequency electromagnetic fields (RF EMFs) has been demonstrated against internationally, on the basis that current scientific insights on the safety of vaccinations or 5G RF EMFs are considered unreliable and untrustworthy. In the Netherlands, the science-based monitoring and modelling of nitrogen deposition in Natura 2000 sites, and derivation of (temporary) background values for PFAS in Dutch soil are the latest of a long list of issues where scientific methodologies, data or conclusions were openly disputed, dismissed or demonstrated against. Meanwhile, the availability of contrasting, or even mutually exclusive strings of scientific evidence allow stakeholders and policy makers to substantiate similarly contrasting policy preferences. In return, scientific advisory institutes, such as the Dutch National Institute for Public Health and the Environment (RIVM), are seeking ways to respond to these trends.*

This thesis aims to study reasoning and argumentation underlying international differences in stakeholders' preferences for regulation, taking endocrine disrupting substances (EDSs) as an example. The issue of EDSs has been likened to Pandora's box (Klinke and Renn, 2002), on the basis of uncertain probabilities of occurrence of effects ('will something bad happen?'), uncertainties in the extent of damage ('how bad will it then be?') and the high persistency of these potential impacts ('when something bad happens, it will remain so for a long time). This study is performed in the context of the ongoing scientific controversy over EDSs hazards and risks, the changing role of scientific knowledge and advice in society and governmental policy making processes, and the associated challenges for the work of scientific advisory institutes. Specific attention will go out to assessing the range of values present in the scientific discourse on EDSs, and argumentation theory will be used to make explicit implicit premises contained in the argumentation under scrutiny, including normative elements.

1.1 ENDOCRINE DISRUPTING SUBSTANCES: A MODERN ENVIRONMENTAL HEALTH RISK ISSUE

In the modern age, chemical substances have become ingrained in human life. A wide variety of substances has been developed that have led to a wide range of beneficial products. To be able to produce durable though flexible plastic products, such as flooring, tubing or food packaging, phthalate-type plasticizer additives have long been applied. Various types of food are canned to increase their preservability, and an inner lining of plastic prevents corrosion of the can by its contents. To ensure optimal yields in the (intensive) farming of cereal grain crops, fungicides, such as azole-class substances, are frequently applied to control crop diseases. However, in recent decades, the safety of these and other products and substances has become subject of intense discussion. Plastics may contain various phthalates (e.g. DEHP) and Bisphenol A, which have endocrine activity. Several azole-class fungicides have also been targeted as having endocrine disrupting properties. Beyond these examples, there is a multitude of compounds and compound classes that have been implicated as possible endocrine disrupting substances (see e.g. EP, 2019; Kortenkamp et al., 2011; WHO-UNEP, 2013).

'Endocrine disruptor' is a popularly used denominator for a substance with the characteristic to 'mimic' the activity of (endogenous) hormones, thereby influencing or even disrupting the endocrine system of humans and wildlife. Other names used to refer to a substance with endocrine disrupting properties are 'endocrine disrupting chemical' (EDC), 'endocrine active substance' (EAS) or 'endocrine disrupting substance' (EDS). Throughout this thesis, the latter term 'endocrine disrupting substance' will be used, following the terminology in relevant European regulatory frameworks, such as the Plant Protection Products Regulation (PPPR) or Biocidal Products Regulation (BPR). These frameworks consistently refer to "substances … [having] endocrine disrupting properties', rather than using the term 'chemical'.

One of the most well-known, and most notorious examples of an endocrine disrupting substance is the synthetic estrogen diethylstilbestrol (DES). From 1938 (i.e. the year of its synthesis) until the 1970's, DES was extensively prescribed to pregnant women to prevent miscarriages (CDC, 2020). In 1971, Herbst and colleagues were the first to identify the association of *in utero* exposure to DES with the manifestation of vaginal clear cell adenocarcinoma in the, then, teen-aged daughters of the DES-exposed mothers. The application of DES as an estrogenic pharmaceutical, to attempt to remedy purported estrogen deficiencies in pregnant women, was eventually linked to the disruption of the unborn's estrogen household. Subsequent carcinogenic and teratogenic malformations predominantly manifested in female sex organs (i.e., vagina, uterus), since these are the most estrogen sensitive organs. The tragic DES-episode sparked intense scientific interest in the potential of man-made substances to negatively impact the endocrine system of humans. However, endocrine disruption substances (EDS) were not only a concern for humans. Already from the 1940's, evidence was mounting that adverse health effects in specific wild-life species were caused by exposures to EDS in the environment (Matthiessen, 2003). A well-known example is the enormous decline of the alligator population in Lake Apopka, Florida, in the beginning of the 1980's, following a chemical spill containing high concentrations of dicofol and DDT, including

its metabolites (Guillette et al., 1994). The highly limited reproductive success of the alligator population was found to be a consequence of, among others, the abnormal reproductive development of male and female juvenile alligators, which, in turn, was attributed to exposure to estrogen active contaminants.

Though concerns about the impact of human and environmental EDSs had been growing for decades, it was not until 1991 when the term 'endocrine disruptor' was coined by the participants of the Wingspread Conference (Markey et al, 2003, see also the consensus statement in Colborn and Clement, 1992). One of the key initiators for the conference was Theo Colborn, who studied developmental and reproductive abnormalities found in wildlife living in the Great Lakes region. Her efforts to link these observations to exposure to environmental pollutants, and the subsequent development of the endocrine disruption hypothesis are described in the popular book 'Our Stolen Future' (Colborn, Dumanoski and Myers, 1996). As Rachel Carson's 1962 book 'Silent Spring' did for the risks of indiscriminate pesticide use, 'Our Stolen Future' permanently placed endocrine disruption on the scientific, public and regulatory agenda. Since its release, thousands of scientific publications have been published related to endocrine disruption, worrying items in the media have repeatedly caused public concern about often ubiquitous EDSs (e.g. Bisphenol A used in baby bottles and food packaging materials) and lawmakers around the world have enacted, or are attempting to introduce, a range of regulations to deal with these (potential) risks. Accordingly, the endocrine disruption hypothesis itself is universally accepted across scientific experts and regulators, though it remains debated which substances should in fact be categorized as substances with endocrine disrupting properties, and at which levels they pose risks to the health of humans and the environment.

There are substances, such as DES, that became clear-cut, widely accepted cases of EDS. For many chemical substances, of which Bisphenol A is a prime example, its EDS status remains disputed. In fact, the enormous scientific attention has polarized the scientific debate, with some experts pointing to scientific evidence that proves the endocrine disruptive properties of a certain substance, and other experts pointing to different data exonerating the substance of these claims.

In the absence of (absolute) certainty that key high-volume chemicals like Bisphenol A, phthalates and certain pesticides are established EDSs (i.e., on the basis of their intrinsic hazardous properties or additional considerations related to risk assessment, such as substance potency or actual levels of exposure), it may not be surprising that there are significant differences between countries as to how they think EDSs should be regulated. Such decisions depend on how a substance's beneficial aspects are weighted against the (inconclusive) scientific data available pointing towards endocrine disruption-mediated adverse effects. For example, the question is how higher crop yields, subsequent increased food security and increased farmer profits due to the use of pesticides are valued against the available toxicological, epidemiological and endocrinological weight-of-evidence. The same holds for evaluation of this evidence vis-à-vis profits made, employment associated with and innovations enabled by the production and processing of plastics.

1.2 ENDOCRINE DISRUPTING SUBSTANCES: A SCIENTIFIC PERSPECTIVE

In 2002, the first global assessment of the state-of-the-science report of endocrine disruptors was published by the WHO/IPCS. The definition proposed in the report, that is, '*an endocrine disruptor is an exogenous substance or mixture that alters function(s) of the endocrine system and consequently causes adverse health effects in an intact organism, or its progeny, or (sub)populations*' (WHO-IPCS, 2002), has become widely considered as an accurate starting point for the scientific study of EDSs. Its strength lies in the explicitly required relationship between three elements of an EDS: a.) its capacity to induce adverse health effects (e.g. breast cancer), b.) the occurrence of an endocrine disrupting mechanism of action (e.g. disruption of the estrogen household), and c.) a causal relationship between the endocrine mechanism of action and the observed adverse effects in an intact organism (i.e., the observed breast cancer is caused by disruption of the estrogen household following exposure to an exogenous substance). However, intense discussions have been taking place about what are the 'right' interpretations of the principles and concepts alluded to by (the different elements of) EDS definition (contrast e.g. Nohynek et al., 2013, Lamb et al., 2014, and Autrup et al., 2015 with Vandenberg et al., 2012, Zoeller et al., 2015 and Bergman et al, 2015). Four main controversies are summarized here.

First, the *issue of homeostasis* has been subject of debate. Some experts feel that the endocrine system strives to maintain homeostasis in the body, by being able to compensate for external stressors. Thus, up to a degree, responses to (external) stressors can be compensated within homeostatic control and do not result in adverse health effects, whereas beyond such threshold the system is disrupted resulting in adverse effects (see e.g. Piersma et al, 2011). Others argue that any exposure will lead to a response of, and thus adversely affect, the endocrine system that is considered crucial to development, growth and reproduction. Such response in itself should be considered adverse, and therefore no threshold of adversity exists.

Second, different perspectives exist towards the *issue of life stage specificity of EDS effects*. Some experts stress that life stage dependent differences in the endocrine system affect the sensitivity to exposures. In general terms, this means that the life stage in which exposure occurs influences the occurrence and severity of adverse health effects. For instance, there is discussion around the question of to what extent the embryo-fetal compartment is protected by maternal metabolism and excretion of xenobiotics (see e.g. Kortenkamp et al., 2011).

Third, different perspectives exist towards the *issue of dose-response interpretation*. The traditionally assumed dose-response of toxic compounds in toxicology is a monotonic sigmoid curve on the logarithmic scale, assuming that at low doses no adverse effects occur. Whether such threshold actually exists in general in a diverse human population (with extreme differences in sensitivity) is a matter of debate (see e.g. Varret et al., 2018; Vandenberg et al., 2012). There is also extensive debate about whether adverse endocrine mediated effects may arise at low doses and whether non-monotonic dose response curves exist. The latter implies that although at certain doses no adverse effects may occur, lower doses may nevertheless cause

adverse health effects. This situation would affect the determination of Acceptable Daily Intake levels, as no safe exposure level can then be determined under such conditions.

Fourth, the *issue of compound potency* is debated, as there is discussion of whether compound potency should be included in the categorization of a compound as an EDS (see e.g. Borgert, Matthews and Baker, 2018). Many compounds display some endocrine activity in isolated *in vitro* assays at relatively high concentrations only. The relevance of such findings in view of actual exposure levels occurring in the intact organism is a matter of debate.

Besides discussions about the interpretation of key principles and concepts related to EDSs, there also appear to be disagreements between experts about other issues. Key examples are how and when causality can be established between (ubiquitous) exposure to EDSs and (multifactorial) adverse effects and diseases, how weight-of-evidence evaluations should be undertaken and how the (extensive) scientific literature on EDSs should be systematically and accurately assessed.

1.3 ENDOCRINE DISRUPTIVE SUBSTANCES: A REGULATORY PERSPECTIVE

The regulation of substances that are associated with having endocrine disrupting properties is highly complex. In the EU, chemical substances are regulated on the basis of the most sensitive adverse outcomes following exposure, such as skin irritation or specific organ toxicity. Current testing strategies adopted in the regulatory assessment of chemical substances have primarily been developed to show a range of such adverse outcomes. However, the unique regulatory feature of EDSs is that their categorization is primarily based on mechanism of action. Animal studies, which make up an important part of the suite of regulatory testing strategies, have not been developed to shed light on mechanisms of action. Furthermore, EDSs can cause different effects than currently surveyed in current regulatory testing (Graven et al., 2016). As current regulatory studies do not provide the means to sufficiently elucidate endocrine mechanisms of action and lack certain endocrine disruption endpoints, the regulation of EDSs requires a different approach to both hazard and risk assessment methodologies as well as to chemical substance regulation in general. In addition, a wide variety of different substances that are regulated by different regulatory frameworks have been associated with endocrine disruption, and there are differences as to the provisions on EDSs in the various EU regulatory frameworks. Table 1.1 summarizes some examples of these issues.

From Table 1.1, it also becomes clear that a uniform regulatory approach for EDSs in the EU is currently lacking. For several years, the European Commission has been working to establish such cross-regulation criteria, based on the state of the science, for the management of EDSs. The Roadmap of the Commission (EC, 2014) provides four possible sets of criteria that could be implemented for the regulatory identification of EDSs. As part of this Roadmap, an impact assessment was performed, which also included a public consultation. A wide range of EU member state governments, non-EU governmental trading partners and other stakeholder organizations responded to this consultation.

From the report on the results of the consultation (EC, 2015a), which presents the highlights of every individual response, it becomes clear that there is no unanimously shared perspective among EU member states about what the ultimate regulatory criteria should entail. Rather, different EU member states point towards different regulatory strategies. For example, a major area of debate is whether potency should be included in the EU regulatory criteria. In the responses on behalf of several EU member states, such as those of the Scandinavian countries (including EAA country Norway) and France, a regulatory option is favored that explicitly omits the inclusion of potency. Alternatively, in the response on behalf of the UK government, a regulatory option that does include potency is favored.

Table 1.1: The EU regulatory frameworks relevant to EDSs, including the phrase that indicates the strength of evidence for establishing causality. This table is adapted from EC, 2014. *These examples are meant to give the reader unfamiliar with EDSs a general idea about the types of substances that have been or are currently associated with having endocrine disrupting properties, and how these substances would be subjected to different strength-of-evidence requirements. These examples should by no means be understood as designations or exonerations of these substances being EDSs by the author.

EU Regulatory framework	**Type of chemical substances**	**Example of substance/application discussed in the context* of endocrine disruption**	**Phrase indicating strength of evidence to establish causality**
Registration, Evaluation, Authorization and Restriction of Chemicals	Industrial chemicals	Bisphenol A	"*for which there is scientific evidence of probable*"
Medical Devices Directive	Medical Devices	Medical tubing containing phthalate plastic softeners	"*for which there is scientific evidence of probable*"
Plant Protection Products Regulation	Pesticides (used to protect crops from e.g. harmful organisms or weeds)	Glyphosate (main active ingredient of weed killer 'Roundup')	"*which may cause*"
Biocidal Products Regulation	Biocides (used to combat harmful organisms in other settings than agri- and horticulture)	Triclosan (antimicrobial agent used in consumer products)	"*which may cause*"
Water Framework Directive	Chemicals present in aquatic environments	Atrazine (weed killing pesticide) found in surface water	None; the Water Framework Directive does not contain specific provisions for the identification of EDSs
Cosmetics Regulation	Cosmetics	Parabens (preservative ingredients in e.g. personal hygiene products)	None; the Cosmetics Regulation does not contain specific provisions for the identification of EDSs

The notion that EU member states have different preferences for EDS regulation is further supported by observed differences in how individual member states prefer to regulate EDSs. With regard to Bisphenol A, the French government issued a ban of BPA-containing products coming into direct contact with food for young children and babies in 2010, which was subsequently adopted EU-wide in 2011. In 2014, a restriction proposal for the use of BPA in thermal paper was submitted by France, which was taken up by the European Commission in 2016, on the basis of risks occurring to workers. In 2015, the French government banned all BPA-containing food packaging materials, effectively in defiance with existing relevant EU regulations. In documentation of DG Grow of the European Commission, the French measure is referred to as 'fully disproportionate', signaling the discomfort of certain departments of the Commission with the ban (EC, 2015b). The Swedish government banned the use of BPA in water pipe lining in 2016, though this application of BPA remains unrestricted in the rest of the EU. Another example is provided by the Danish action on the chemical UV filter 4-MBC. In 2001, the Danish Environmental Protection Agency effectively removed sunscreens containing 4-MBC from the Danish market by making a voluntary agreement with manufacturers and importers that 4-MBC was banned in products for children under 12 years. This intervention was a response to preliminary results from exploratory toxicity tests (i.e. Schlumpf et al., 2001) that were considered sufficiently alarming. Using the same data, the Dutch government decided that such an intervention would be unwarranted.

1.4 THE ROLE OF SCIENTIFIC ADVISORY INSTITUTES IN CONTEMPORARY SOCIETY AND POLICY

For scientific advisory institutes, the previous decades have brought significant challenges as to the credibility of the knowledge they produce and share. Traditionally, their 'science-based' advice was typically considered reliable and legitimate by default, since being a knowledgeable scientific institute with high quality expertise in relevant fields was the most important ingredient to being considered a trustworthy authority by policy makers and society. However, this 'credibility by default' appears to have been slowly replaced by a sort of 'wariness by default'.

The Dutch National Institute for Public Health and the Environment (RIVM), providing science-informed advice to various levels of government (local, regional, national, international) about public health and environment issues, feels the need to deal with these societal trends. RIVM is increasingly considered an actor in various scientific and societal controversies. For example, RIVM's roles in monitoring and modelling of nitrogen deposition in Natura 2000 sites, and deriving (temporary) background values for PFAS in Dutch soil have been subject to intense debates. In these situations, it is not uncommon that experts or societal groups openly challenge RIVM's research methodologies, data or conclusions, occasionally leading to public protest gatherings in front of the institute and personal threats to experts.

A wide range of scientific ideas and concepts have been developed to assess why the credibility of scientific advisory institutes like RIVM is increasingly under pressure, and what could be done to reaffirm their position in society. Some of these ideas and concepts are discussed below. What they have in common is their diagnosis that the 'credibility by default' adage was

effectively bound to erode, and that 'wariness by default' is a much more appropriate attitude to scientific knowledge, particularly when the issue at stake is complex, uncertain or subject to normative value differences.

Funtowicz and Ravetz studied how the quality of scientific inputs to policy processes is maintained and coined the term Post-Normal Science, to emphasize that high decision stakes and high systems uncertainties require different approaches for quality assurance then the more 'straight-forward' types of scientific work, such as 'Normal Science' (see e.g. Funtowicz and Ravetz, 1993; Ravetz, 1999). They argue that quality control for Post-Normal Science should include 'extended peer communities', meaning that all those interested and affected should be able to provide their critical evaluations in open dialogue settings.

In their assessments of the way in which scientific knowledge is generated, Gibbons et al. (1994) developed the idea of two modes of science: Mode 1 and Mode 2 science. Characteristics of Mode 1 science, the traditional type of knowledge generation, are that knowledge production occurs predominantly in the academic domain, in accordance with the established division of discrete disciplines. It is primarily driven by the interests of the scientific community, not necessarily by societal interests or the applicability of the resulting knowledge. Alternatively, Mode 2 science is characterized by the generation of knowledge in conjunction with the relevant social actors, driven by their needs and concerns. This type of knowledge production occurs in a transdisciplinary setting and is attentive to the context in which the knowledge is produced and ultimately applied. Accordingly, knowledge produced in Mode 2 science, referred to as 'socially robust knowledge' may be less likely to become controversial than 'reliable knowledge', the product of Mode 1 science (Nowotny, Scott and Gibbons, 2001; Gibbons, 1999).

Jasanoff (2003) argues that 'technologies of hubris' have long been cultivated by researchers and policy-makers, while 'technologies of humility' should be aspired. The former, of which Jasanoff considers traditional 'risk assessment' an example, focus on prediction and control and acquire their power by making claims of objectivity, completeness and analytical rigor. Alternatively, in the complementary idea of 'technologies of humility', technological developments should explicitly be considered in their broad societal context, which necessitates deliberations with interested and affected parties (e.g. citizens, NGO's, industry, policy-makers) about the social, ethical, political and other relevant implications during all stages of development.

From these assessments of the science system, it becomes clear that scientific experts, particularly those working in-between science, policy and society, should explicitly interact with interested and affected parties, and incorporate the knowledge and values of these parties in their scientific research and advisory activities. The necessity for interdisciplinary approaches to scientific research is also a recurring theme. However, meaningful interaction with societal actors, and among experts themselves requires transparency about the personal and professional biases of the interacting experts. There is increasing attention for the idea that scientific research is not free of normative values (see e.g. Douglas, 2000; Elliott, 2017; Sarewitz, 2004). More specifically, science is often credited for its reliance on academic values, such as

independence, meticulousness and skepticism. Apart from the idea that scientific experts will attribute different priorities to different types of academic values in their work, it is also asserted that broader normative values play an important, though often implicit role in scientific research. Examples of such values are human and environmental health and integrity, economic wellbeing and prosperity, equality, freedom, tradition and compliance to ethical standards.

In the risk governance framework of the International Risk Governance Council (IRGC), developed to provide guidance for an inclusive approach to the assessment and management of modern risks, specific attention goes out to value differences among scientific experts. The IRGC distinguishes two types of ambiguities: interpretative and normative ambiguity (IRGC, 2005). Interpretative ambiguity refers to the existence of different interpretations of pieces of scientific evidence, for example about the value and meaning attributed to toxicological evidence as compared to that of epidemiological evidence, or weight-of-evidence. Normative ambiguity refers to differences concerning normative values and (ethical) norms, for example about evaluations of risk acceptability or tolerability. Accordingly, the notion of "more research is needed" could be appropriate for interpretative ambiguity, while settling normative ambiguity requires multi-stakeholder approaches (Hisschemöller and Hoppe, 1995; Renn, 2008). Another notable consequence of normative ambiguity is that it matters considerably which expert provides advice to policy makers.

Various researchers studying the role of scientific experts as policy advisors point to the relevance of normative values in scientific advisory work. Pielke (2007) distinguishes four ideal-typical expert roles (Pure Scientist, Science Arbiter, Issue Advocate and Honest Broker), that differ in relation to the degree of participation within policy processes (ranging from none at all to active participation) and the breadth of an expert's policy advice (advocacy of a particular measure or providing a range of policy alternatives). The appropriateness of an expert advisory role depends on the decision context, characterized by the uncertainty and values consensus surrounding the risk.

Weiss (2003) distinguishes five ideal-typical expert roles (Scientific Absolutist, Technological Optimist, Environmental Centrist, Cautious Environmentalist and Environmental Absolutist) that relate to the level of (precautionary) action advocated, given a certain degree of scientific (un)certainty. The Weiss typologies implicitly relate to normative positions about, for example, the importance of (uncertain) scientific knowledge in policy making as compared to other types of information, the vulnerability of environmental or human physiological systems, and the legitimacy of the Precautionary Principle as a guiding legislative principle in a certain situation.

In an interdisciplinary review of the available literature on the roles of experts as policy advisors on complex issues, Spruijt et al. (2014) identified six factors that influence expert roles. Besides the core values of an expert, the other five factors are: the type of issue, the organization in which the expert works, the type of knowledge of an expert, the changing beliefs of experts and the wider context of the expert. Empirical research of Spruijt and colleagues, addressing the roles of experts in the fields of electromagnetic fields, particulate matter and antimicrobial

resistance, has shown that advisory roles can vary strongly among experts of the same subject matter, though the identified expert roles did not precisely fit the Pielke and Weiss typologies (Spruijt, 2016). Also differences between fields of subject matter were observed, suggesting cultural and/or contextual influences.

The issue of EDSs provides an ideal case to study the challenges that RIVM and similar research institutes face in contemporary society, given the combination of ongoing scientific controversy and the divergent regulatory preferences of EU member states.

1.5 RESEARCH QUESTIONS AND METHODOLOGY

In the research project outlined in this thesis, the main aim is to study reasoning and argumentation underlying international differences in preferences for EDS regulation, in the context of the ongoing scientific controversy over EDS hazards and risks. The main research question is: *What argumentation underlies the international differences in preferences of stakeholders towards the regulation of EDSs?*

This question is addressed in three separate studies; one literature review and two empirical studies. In preparation for and partly parallel to the literature review and the first empirical study, an exploratory case study using semi-structured interviews was performed to better gauge the scope and dimensions of an actual discourse. This case study is summarized in paragraph 1.5.2. In the two empirical studies, we used Pragma-Dialectical Argumentation Theory to perform argumentation analysis of the scientific controversy over EDS and of different stated preferences of government entities and stakeholder organizations. This theory of argumentation allows, to the degree possible, an impartial and systematic analysis and interpretation of the arguments put forward by the various parties to support their expressed standpoints.

1.5.1 Pragma-Dialectical Argumentation Theory

In order to systematically perform an argumentative analysis, we use approaches and insights provided by the pragma-dialectical argumentation theory, developed by van Eemeren and colleagues (van Eemeren et al. 2014; van Eemeren and Grootendorst, 1984). According to PDAT, the goal of an argumentative discussion is solving a difference of opinion by means of acceptable argumentation. The name of PDAT is inspired by its intellectual foundation; the study of pragmatics and the study of dialectics. Scholars in the field of pragmatics study how 'real life' communication and interactions attain meaning within a certain context (van Eemeren et al., 2014). For example, the phrase 'I can't get this computer working' could be interpreted in multiple ways. The speaker could simply assert that he cannot get the computer to work. Alternatively, the speaker could also implicitly request help in getting the computer to work. The exact intended meaning of the phrase may then depend on the specific situation, identity of the speaker, intonation of the utterance, body language, prior conversations and other contextual factors. Second, PDAT's dialectical orientation refers to solving differences of opinion in a reasonable manner (van Eemeren et al., 2014). In terms of the PDAT, reasonableness refers to advancing sound argumentative moves that do not impede resolving a

difference of opinion. A typical example of unreasonable argumentation is putting forward argumentation that is based on an invalid reasoning scheme. In paragraph A3.1 of the appendix of Chapter 3, PDAT is discussed in greater detail.

1.5.2 Exploratory case study on BPA in thermal paper using semi-structured interviews

In a first exploration of the depth and breadth of the discourse on EDS regulation, we used semi-structured interviews to study why France drafted a proposal to restrict the use of Bisphenol A (BPA) in thermal paper (i.e. paper used in cash receipts). This case study allowed us, on the one hand, to study the argumentation underlying the French efforts to specifically target thermal paper, and gain insight into the roles and potential influence of the wide range of stakeholders, including the general public, in the societal and political debates on this issue. This explicitly includes any argumentation referring to scientific knowledge and principles, as this is also a focus of the other chapters of this thesis. On the other hand, we wanted to study the same dynamics in another EU member state with a less pronounced approach to BPA, and contrast these findings to those found for the French situation, to learn what factors contribute to the observed divergent international regulatory preferences for EDSs in general, and BPA specifically (see also Chapter 4). However, due to difficulties encountered during this case study we chose to limit the scope to the French situation and to not replicate it in another EU member state.

We developed a protocol for the semi-structured interviews (see Appendix A1.1). The protocol comprised four parts, each consisting of one or more questions: 1) background questions about the interviewee (e.g. 'could you describe your position with organization X?', 2) perspective towards the science of BPA ('to what degree would you consider the scientific knowledge regarding BPA risks certain, or contested?'), 3) the restriction of BPA in thermal paper in details (e.g. 'could you say something about the stakeholders that have been involved in the process?' and 4) final remarks and follow-up (e.g. 'could you suggest a few people that may be relevant to talk to as well?). We performed one test interview to further fine tune the protocol, and slight changes were made during the course of the study. We subsequently selected eligible interviewees from a variety of French stakeholder, and were able to conduct interviews with experts from ANSES (French Agency for Food, Environment and Occupational Health and Safety) and INERIS (French National Institute for Industrial Environment and Risks), representatives from NGO's WECF (Women in Europe for a Common Future) and RES (French Environmental Health Network), and with a representative from PlasticsEurope, the European Trade Union for the Plastics Industry. These interviews were all held over the phone. One NGO and one French supermarket chain did not reply to our (repeated) request for an interview, the UIC (French Chemical Industry Association) and an individual BPA producer referred us to PlasticsEurope, and a producer of thermal paper and the French trade association for supermarkets kindly declined to be interviewed.

From the interviews, we learned that France adopted a National Strategy on Endocrine Disruptors (Ministry of Ecology, 2020), which formulates various steps to protect the French

population from exposure to EDSs, of which the restriction of BPA in thermal paper was one of multiple. Interviewees also pointed to the very influential role of the 2013 ANSES risk assessment of BPA, the results of which have been covered extensively by the French media. In addition, the French Minister of Ecology, Ségolène Royal, has been very visible in the EDS dossier, and public perceptions of BPA in France have been particularly negative, as has been shown by public perception surveys. All in all, some NGO's have been considered instrumental in placing the issue of BPA (and EDSs in a broad sense) on the French political agenda.

Notably, one of the two interviewed representatives from NGO's explicitly referred to five points that should be key for the assessment of EDSs risks: 1) the role of exposure during vulnerable periods, 2) low dose effects, 3) cocktail effects, 4) latency between early-life exposure and effects later in life, and effects on the epigenome, and 5) transgenerational effects. Alternatively, the interviewee representing PlasticsEurope considered that it is key to consider all data on BPA in accordance with predefined principles on the assessment of the weight-of-evidence and the attribution of weight to different types of studies, to have an earnest, full-encompassing perspective on the potential effects of BPA.

With regard to the choice for targeting thermal papers, a variety of reasons were mentioned during the interviews: 1) all of the uses of BPA were thoroughly scrutinized in France, 2) BPA in thermal paper is applied as a layer, which is removed more easily (thus potentially leading to exposure) in comparison to applications of BPA (e.g. polycarbonates) where it is used as a monomer 3) there was arguably more certainty on the risks posed by exposure to BPA from thermal paper than for other applications 4) thermal paper was a non-food application, meaning it was regulated by REACH and allowed for restriction under the conditions provided by this regulatory framework and 5) industry was already working on a substitute, and, from an economical perspective, a restriction would have relatively limited impact.

On the basis of these results, we argue that (implicit) advocacy coalitions could be discerned in France. We encountered that stakeholders from industry explicitly referred to one another to discern their specific standpoints and argumentation to us, and NGO's that cooperate to inform the French population on the risks of exposure to EDSs and influence French policy makers. We gained insights in the perspectives and roles of specific stakeholders and ultimately learned that the restriction of BPA in thermal paper was part of a broad movement in France to reduce EDS exposure to consumers, patients and other (vulnerable) groups. We also gained some insight into scientific principles underlying the evaluation of EDS risks of the various 'coalitions', which will receive more attention in Chapter 5, in the context of (similar) findings in Chapters 2 to 4. All in all, these insights provided inspiration for the interpretation of the literature and guided the development of the two empirical studies.

1.5.3 Structure of this thesis

Our research question can be approached from various directions. In a theoretical approach, a variety of conceptual frameworks, from a wide range of scientific disciplines, can be used for reflection, where each of these frameworks will provide a different, complementary piece of the puzzle. Accordingly, **Chapter 2** presents an extensive theoretical review, in which eight conceptual frameworks, from different areas of science are explored. Specific attention went out to the role of (controversial) scientific knowledge in regulatory processes, as this is a key theme in this thesis. Note that this chapter does not exclusively focus on the issue of EDSs (as Chapters 2 and 4 do), but on environmental health risk issues in general, since the conceptual frameworks discussed can also be used to study other issues, such as electromagnetic fields, particulate matter or noise pollution.

In **Chapter 3**, a study piloting the use of PDAT is presented, where two high-profile scientific publications have been selected for analysis (Lamb et al., 2014 and Bergman et al, 2015) on the basis of their contrasting perspectives towards key issues in EDS science. This study was initiated to gain a deeper understanding of the type and nature of key issues in the EDS scientific controversy. This study would then serve as input to further assess the relationship between controversial scientific knowledge and divergent regulatory preferences. We aim to make explicit differences in starting points that appear to play a role in the discussion over EDS science, but often remain implicit. We also assess whether these differences in starting points account to different interpretations of the underlying scientific knowledge (interpretative ambiguity), or whether judgments about broader values and (ethical) norms are involved (normative ambiguity).

The ideas and concepts identified in the theoretical review and experiences gained from our first study with PDAT were then applied to our second empirical study, presented in **Chapter 4**. We performed an argumentation analysis on the responses to the public consultation of the EU, concerning the impact assessment of four option for regulatory criteria to identify EDSs. A wide variety of stakeholders, including national governments, participated in this consultation, thereby providing their preferred regulatory option and supporting argumentation. Our main aim was to study the underlying argumentation provided by different governments and other stakeholders in support of their expressed standpoints, i.e. their preferred options for regulatory criteria to identify EDS in the EU. Accordingly, PDAT was used to identify the range of arguments addressed in all the responses of national-level governmental entities and key stakeholder organizations, and to compare the argumentation of responses that supported differences regulatory options.

Finally, **Chapter 5** contains a general discussion. We will draw some links between our theoretical work and the empirical findings and provide some suggestions on how to proceed in the debates on EDS science, policy and regulation. We will also share some experiences and lessons learned from our application of PDAT. Finally, some implications for future research will be discussed.

Why do countries regulate environmental health risks differently? A theoretical perspective

Sander C.S. Clahsen, Irene van Kamp, Betty C. Hakkert, Theo G. Vermeire, Aldert H. Piersma and Erik Lebret

Published in:
Risk Analysis
Volume 39, Issue 2, 2019

ABSTRACT *Why do countries regulate, or prefer to regulate, environmental health risks such as radiofrequency electromagnetic fields and endocrine disruptors differently? A wide variety of theories, models, and frameworks can be used to help answer this question, though the resulting answer will strongly depend on the theoretical perspective that is applied. In this theoretical review, we will explore eight conceptual frameworks, from different areas of science, which will offer eight different potential explanations as to why international differences occur in environmental health risk management. We are particularly interested in frameworks that could shed light on the role of scientific expertise within risk management processes. The frameworks included in this review are the Risk Assessment Paradigm, research into the roles of experts as policy advisors, the Psychometric Paradigm, the Cultural Theory of Risk, participatory approaches to risk assessment and risk management, the Advocacy Coalition Framework, the Social Amplification of Risk Framework, and Hofstede's Model of National Cultures. We drew from our knowledge and experiences regarding a diverse set of academic disciplines to pragmatically assemble a multidisciplinary set of frameworks. From the ideas and concepts offered by the eight frameworks, we derive pertinent questions to be used in further empirical work and we present an overarching framework to depict the various links that could be drawn between the frameworks.*

2.1 INTRODUCTION

International differences in the management of environmental health risks occur frequently. For example, the United States and the European Union regulate potential risks posed by pesticide contaminants in drinking water differently. In the United States, a maximum contaminant level (MCL) is derived for a single pesticide, based on the available toxicological evidence and an analysis of the costs and health benefits associated to the proposed MCL (EPA, 2016). Alternatively, in the European Union, almost all pesticides approved for use are subject to the same maximally allowable concentration (MAC) of 0.1 μg/L, which is supposed to function as a precautionary-based "surrogate zero" level (Dolan et al. 2013). Similar differences can be observed between European countries. Redmayne (2016) has shown that regulatory approaches toward children's exposure to radiofrequency EMFs differ widely among European countries. In addition, Löfstedt (2011) argues that several European countries have different preferences with regard to the management of the industrial chemicals Bisphenol A and Deca BDE.

In this theoretical review, we will explore various theoretical perspectives that help understand these international differences in risk management. We are interested in studying the reasoning behind countries' risk management strategies, and to better understand why these strategies differ between countries. To further specify our understanding of the term "international differences in risk management strategies," we distinguish two dimensions in which risk management can differ across countries. First, countries can select different risks for regulation. As Wiener et al. (2011) point out, there are key differences between the United States and Europe in terms of the policy areas that are thought to require precautionary regulatory interventions. Second, in the event that countries set out to manage the same risk issue, then differences in the stringency of the ultimate measures can occur. For example, Rothstein et al. (2017) found that there are variations in the stringency of occupational health and safety regulations in Germany, France, the United Kingdom, and the Netherlands. The theoretical frameworks discussed throughout this article could be used, although to different extents, to reflect on these two dimensions.

We are specifically interested in the role of scientific expertise within risk management processes. This interest comes from our working experience as experts in institutes and expert committees with a mission to translate scientific knowledge for environmental and public health policymakers. In this capacity, we have experienced that the relationship between science and policy is not as straightforward as the term "evidence-based policy" seems to suggest. Decades of research into this relationship in the domain of environmental health risks theoretically and empirically underlines these experiences (Funtowicz & Ravetz, 1993; Hoppe, 2009; Jasanoff, 1990; Pielke, 2007; Sarewitz, 2004; Spruijt et al., 2014). The variety of conceptual frameworks discussed in this article shed light on various aspects of complex science–policy interactions.

As a result, a variety of explanations for international differences in risk management strategies are explored: Do the differences stem from different overall interpretations by experts of the

underlying scientific evidence, that is, interpretative ambiguity (after Renn, 2008)? Do these differences occur because experts hold different views about the acceptability of risks, that is, normative ambiguity? Or does expert judgment only play a minor role in the final judgment and do differences emerge through influences from other stakeholders? Or do the differences follow from differences in (national) cultures? The outcomes of this review will be used to structure stakeholder interviews and document analyses used in further case studies that aim to study international differences in environmental health risk management.

Overall, the aim of the present review is to provide a theoretical foundation for empirical research that seeks to understand why countries frequently manage environmental health risks differently, and how expert policy advisors specifically, and scientific knowledge in general, are involved in the development of these different risk management strategies. To that end, we will discuss eight different conceptual frameworks: the Risk Assessment Paradigm, research into the advisory roles of experts in risk management processes, the Psychometric Paradigm, the Cultural Theory of Risk, several participatory approaches to risk assessment and risk management, the Advocacy Coalition Framework, the Social Amplification of Risk Framework, and Hofstede's Model of National Cultures.

2.2 METHODS

The selection of the eight conceptual frameworks included in this article occurred through a pragmatic rather than a systematic approach. A systematic selection of relevant frameworks would prove challenging because concepts and terminology vary across scientific domains. Through cognitive distance between experts from different domains, different words are used for (nearly) similar concepts and the same words for different concepts (Lebret, 2016). Moreover, in some academic fields, main thoughts and concepts are published in books rather than in peer-reviewed journal articles or proceedings. Alternatively, we drew from various sources to identify relevant conceptual frameworks. First, a variety of theories and models has been brought forward in scientific debates occurring in literature, in conference discussions, in expert committee meetings and discussions occurring between relevant colleagues. Second, we were familiar with several conceptual frameworks through our knowledge and experiences in the fields of environmental health, food safety, regulatory toxicology, risk perception, risk governance, science and technology studies, and (organizational) psychology. Third, we used a "snowball" approach to identify further relevant frameworks.

2.3 RESULTS

2.3.1 Understanding the Concept of "Risk"

Because the concept of "risk" is referred to extensively throughout this article, we think clarity about our specific understanding of this concept is required. Indeed, "risk" can be defined in numerous ways (Aven & Renn, 2010). Accordingly, Rayner (1992) argues that "risk" is a "polythetic" concept, meaning that a wide variety of equally legitimate, though complementary and sometimes contrasting, understandings of the concept of "risk" coexist. In the present article, two influential though contrasting understandings of "risk" play a key role.

First, in the risk assessment sciences (i.e., toxicology, epidemiology, and exposure sciences), risks are interpreted in a quantitative and technical sense. Risks are typically assessed in terms of their physical characteristics, in a "probability of occurrence * magnitude of effects" fashion. The Risk Assessment Paradigm is the key intellectual foundation of this type of risk assessment.

Alternatively, risk perception research (i.e., research focusing on how others than subject-matter experts typically assess risks) has revealed that other considerations than the physical characteristics of a risk often play an important role in risk evaluation. Two theoretical perspectives are dominant here: the Psychometric Paradigm and the Cultural Theory of Risk. Both perspectives refer to the "construction of risk" (although in different ways), meaning that risks are not considered fixed, objective entities interpreted the same by everyone. Rather, individuals and groups construct risks differently, depending on their personal values, social and cultural contexts and other economic, legal, and ethical considerations. On these grounds, one could also assert that technical risk assessment is a specific type of risk perception, developed, systematized, and maintained by expert risk assessors, in an effort to thoroughly characterize a risk in the most objective way possible.

2.3.2 Overview of Eight Conceptual Frameworks

2.3.2.1 Risk Assessment Paradigm

Policymakers often require the assessment of a particular environmental health risk, such as risks associated with exposure to particulate matter, electromagnetic fields, or specific chemical substances. Such risk assessments are often performed in accordance with the Risk Assessment Paradigm (RAP), as originally outlined in the Red Book of the U.S.-based National Research Council (NRC, 1983). In the IPCS Risk Assessment Terminology report (IPCS, 2004), "risk assessment" is defined as "a process intended to calculate or estimate the risk to a given target organism, system or (sub)population, including the identification of attendant uncertainties, following exposure to a particular agent, taking into account the inherent characteristics of the agent of concern as well as the characteristics of the specific target system" (p. 14). Four steps together form the risk assessment process: hazard identification, dose–response assessment, exposure assessment, and risk characterization; see also Figure 2.1. According to a later report of the NRC, the main achievement of the 1983 version of the paradigm is its popularizing of the distinction of the risk assessment process from the process of risk management (NRC, 1996). Other, later models that describe or prescribe the process of risk assessment, such as the model developed by Covello and Merkhofer (1993), similarly maintain this distinction, signaling the influence of this idea. Nevertheless, in its later works (see NRC, 1996, 2009), the NRC recognized the need for risk assessors, risk managers, and other stakeholders to cooperate at various stages of the risk management process (e.g., the problem formulation stage), blurring the lines between risk assessment and risk management (see also Section 2.2.5).

From the point of view of this paradigm, regulatory differences can occur due to differing interpretations of scientific evidence by experts and different (methodological) choices made by experts in situations of uncertainty. In the environmental health domain, scientific

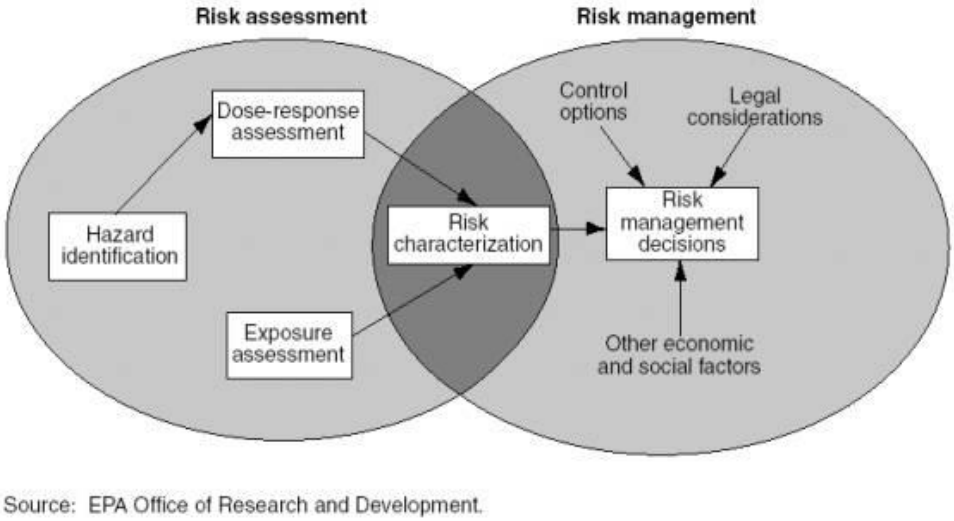

Figure 2.1: The Risk Assessment Paradigm as described in the NRC's 1983 Risk Assessment in the Federal Government: Managing the Process. Source: EPA, 2020. Permission to reprint this figure has been obtained from the copyright holder (U.S. Environmental Protection Agency).

knowledge pointing to a "high-risk" situation can often coexist with knowledge suggesting a "low-risk" situation. For example, experts disagree about the carcinogenic properties of glyphosate and, accordingly, about the risks associated with exposure to this herbicide (compare FAO/WHO, 2016 and EFSA, 2015 with IARC, 2017). In addition, Beronius, Rudén, Hakansson, and Hanberg (2010) note that the outcomes of the various risk assessments of Bisphenol A differ considerably, illustrating the apparent ambiguity over whether or not BPA is shown to pose a risk to (parts of) the population. The phenomenon that experts interpret certain scientific evidence differently is also known as *interpretative ambiguity* (Renn, 2008). As a result of this ambiguity, the degree to which either line of evidence is represented within a country may differ internationally; experts in country A may be pointing more to evidence suggesting a "high-risk" situation, whereas experts in country B mostly refer to evidence suggesting a "low-risk" situation when providing policy advice. In short, when studying international differences in risk management strategies, one could investigate which specific lines of (risk assessment) evidence were used to support the risk management process, and how the evidence referred to within these processes differs from country to country.

2.3.2.2 Expert Advisory Roles in Risk Management Processes

When highly specialized, in-depth knowledge is required to thoroughly understand the hazardousness of a risk, policymakers can turn to experts for science-based policy advice. Traditionally, policymakers and experts should interact while observing a clear division of labor: policymakers are responsible for the normative process of policy development, whereas experts are responsible for the independent, value-free process of scientific knowledge generation. The (conceptual) separation between risk assessment and risk management exemplifies this idea.

However, this traditional separation of roles does not accord with the complex relationship between experts and policymakers as observed in reality (see e.g., Jasanoff, 1990). The roles of experts in policy processes may particularly become ambiguous when risks are complex, uncertain, and contested, like the risks associated with exposure to RF EMFs (Spruijt et al., 2014).

We have identified several theoretical typologies of expert roles that describe the particular options (theoretically) available to experts when acting as a policy advisor. Pielke (2007) distinguishes four ideal-typical expert roles (see also Figure 2.2). These expert roles pertain to the degree of participation within policy processes (ranging from none at all to active participation) and the breadth of an expert's policy advice (advocacy of a particular measure or providing a range of policy alternatives). The *Pure Scientist* typically performs research out of intrinsic interest and does not interact with policymakers in any way whatsoever, whereas the *Science Arbiter* answers factual questions posed by policymakers. The *Issue Advocate* advocates the implementation of one specific policy option, thus reducing the scope of possible policy options, whereas the *Honest Broker* aims to provide an overview of all scientifically legitimate policy alternatives, thus broadening the scope of possible policy options. Pielke (2007) proposes that all four roles are equally legitimate, but the appropriateness of a specific role depends on the decision context, characterized by the uncertainty and values consensus surrounding the risk.

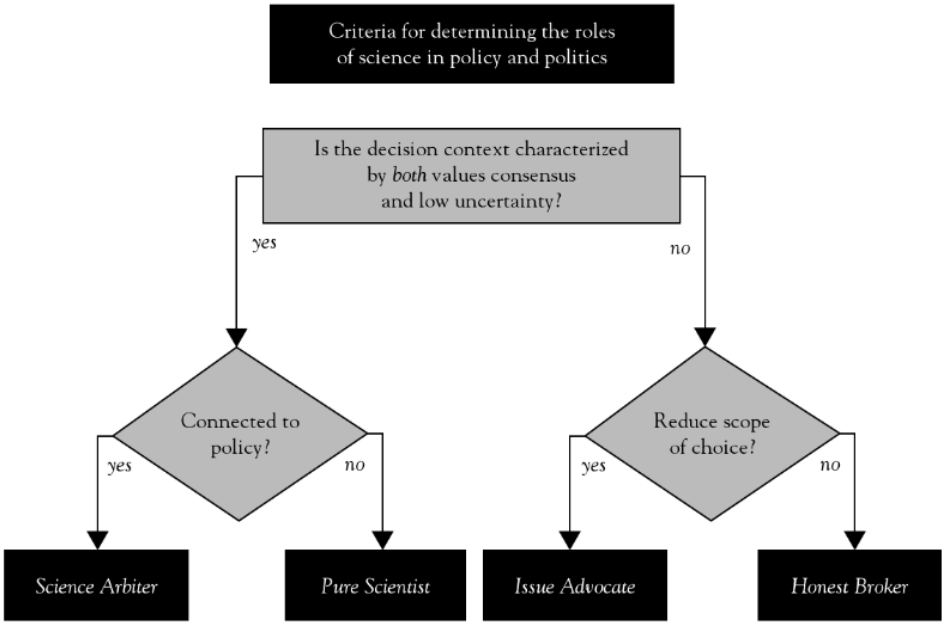

Figure 2.2: Four ideal typical roles developed by Pielke: Science Arbiter, Pure Scientist, Issue Advocate an Honest Broker. Source: Pielke, 2007. Permission to reprint this figure is automatically granted by the copyright holder (Cambridge University Press) under condition of acknowledgment.

Furthermore, Weiss (2003) distinguishes five ideal-typical expert roles (see also Figure 2.3). These expert roles pertain to the technical contents of policy advice, particularly in relation to the level of (precautionary) action advocated, given a certain degree of scientific (un)certainty. The *Scientific Absolutist*, typically advocating "science-based regulation," requires high levels of scientific certainty before supporting "measures against the most serious aspects" of a technological development that may pose potential dangers to the environment. Alternatively, the *Environmental Absolutist*, typically advocating relatively early precautionary action, requires much less scientific certainty before advising the same measure, based on the norm that man should inherently protect the environment from potential (man-made) dangers. The *Technological Optimist*, *Environmental Centrist*, and *Cautious Environmentalist* hold intermediate positions.

In an interdisciplinary review of the available literature on the roles of experts as policy advisors on complex issues, Spruijt et al. (2014) identified six factors that influence expert roles. These are: the type of issue, the type of knowledge of an expert, the core values of an expert, the organization in which the expert works, the wider context of the expert, and the changing

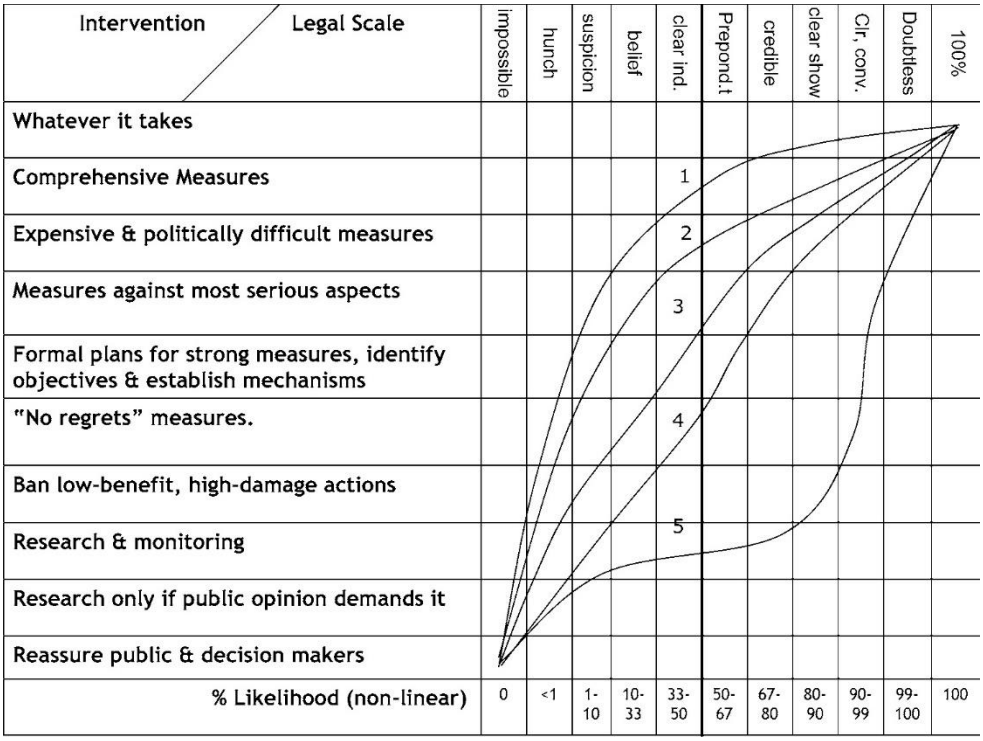

Figure 2.3: The level of intervention to address severe hazards to the environment is plotted against the level of scientific (un)certainty and five roles of experts. The probability scale is asymmetrical and nonlinear. The curves correspond to the following five expert roles: 1. Environmental Absolutist, 2. Cautious Environmentalist, 3. Environmental Centrist, 4. Technological Optimist, 5. Scientific Absolutist. *Source:* Weiss, 2003. Permission to reprint this figure has been obtained from the copyright holder (Springer Nature).

beliefs of experts. Though this review included a wide variety of theoretical ideas from various scientific disciplines, Spruijt, Knol, Torenvlied and Lebret (2013) acknowledge that empirical evidence supporting this theory is scarce. To fill this gap, Spruijt and colleagues studied the advisory roles of experts in the fields of EMF (Spruijt, Knol, Petersen, & Lebret, 2015), particulate matter (PM) (Spruijt, Knol, Petersen & Lebret, 2016), and antimicrobial resistance (Spruijt, 2016). They found that advisory roles can vary strongly among experts of the same subject matter, though the identified expert roles did not precisely fit the Pielke and Weiss typologies. In short, participating EMF and PM experts had different judgments as to the necessity of additional (precautionary) measures in EMF risk management (Spruijt et al., 2015) and PM risk management (Spruijt et al., 2016). Furthermore, PM experts disagreed about whether it is their responsibility to act as an "issue advocate," that is, recommend the policy option that the expert deems most suitable (Spruijt et al., 2016).

From the theoretical and empirical work on the roles of experts as policy advisors, we draw the main message that science–policy interactions are highly intricate by nature. A recurring theme is the role of normative values held by experts and how such values could (or should) influence their advisory roles. For example, the five expert roles developed by Weiss implicitly relate to normative positions about the importance of scientific knowledge in policy making, the vulnerability of environmental or human physiological systems, and the legitimacy of the Precautionary Principle as a guiding legislative principle. In turn, these expert roles draw attention to the possible variations in stringency of (precautionary) measures advised by expert policy advisors. When experts hold different value-based positions with regard to the acceptability of a risk, this can be referred to as normative ambiguity (in contrast to interpretative ambiguity) (Renn, 2008). A key consequence of normative ambiguity in science is that it matters considerably which expert provides advice to policymakers. The idea that different experts may provide different policy advice could be used as an explanation for observed international differences in risk management.

2.3.2.3 Psychometric Paradigm

In essence, research within the Psychometric Paradigm (PP) aims to unveil factors that determine the risk perception of people (Siegrist, Keller, & Kiers, 2005). This type of research is quantitative by nature, employing questionnaires to ask respondents about the perceived riskiness, acceptability, and the desired level of risk regulation for a wide variety of human activities and technologies (Slovic, 1987). Notably, these findings are then compared to the respondents' scorings on qualitative characteristics of risk that are hypothesized to influence risk perceptions. For example, Fischhoff, Slovic, Lichtenstein, Read, and Combs (1978) asked respondents to rate each of the 30 activities or technologies included in the study on nine characteristics of risk. These were: voluntariness of risk, immediacy of effect, knowledge about risk (exposed), knowledge about risk (science), control over risk, newness, chronic–catastrophic, common–dread, and severity of consequences. Research into the statistical relationships between these different characteristics of risk, by means of factor analysis, showed that some characteristics correlated highly with one another. In fact, three higher-order factors have been discerned repeatedly: "dread risk," "unknown risk," and "number of people exposed"

(Slovic, 1987; Slovic, Fischhoff, & Lichtenstein, 1976). The specific characteristics of risk that define these higher-order factors can be observed in Figure 2.4. Individual risks can then be placed in a two-dimensional factor plot, in accordance with their scoring on the aggregated "dread risk" and "unknown risk" factors. The resulting factor plots, sometimes referred to as "cognitive maps," represent the average risk perceptions of the entire group of respondents (Slovic, 1987).

The PP resulted from a research program aimed at studying cognitive processes that determined man's responses to risks, particularly natural hazards (Slovic, 1987). Attempts to explain why the perceptions of some hazards seemed to defy the outcomes of statistics-based risk assessment (e.g., exemplified by societal resistance to statistically speaking lowrisk nuclear power) focused on the limitations of human cognition. According to Slovic et al. (1976), the idea that any decisionmaker is "boundedly rational" (Simon, 1957) functioned as a key source of inspiration here. In addition, Tversky and Kahneman (1974) showed, through various psychological laboratory studies, that humans often revert to heuristics, or mental shortcuts, to make judgments under uncertainty. For example, the availability heuristic could explain why the perceived probability of an event occurring increases when immediate and powerful cues (such as vivid images) are available (Tversky & Kahneman, 1973).

When applying insights from the PP to the study of regulatory differences between countries, the question immediately arises whether there were significant differences in societal risk perceptions, and whether there were major differences in the salience of the risk issue among the populations of the different countries. One would expect that, in a democratic society, public policymakers are reasonably responsive to the concerns of those confronted with the risky activity when developing a risk management strategy (see e.g., Slovic, Fischhoff, & Lichtenstein, 1982). Differences in countries' general societal risk perceptions could then be an explanation for international differences in risk management.

2.3.2.4 Cultural Theory of Risk

The Cultural Theory of Risk (CTR) sets out to explain how individuals and groups select and interpret dangers (Tansey & O'Riordan, 1999). For example, Douglas and Wildavsky (1983) describe various reasons for why some people are specifically attentive to environmental risks, whereas others focus particularly on risks posed by violent crime. According to these authors, such risk selections (and subsequent risk perceptions) are shaped by the social context, and particularly by the *social organization* of the group in which the individual belongs. This premise is mainly supported by insights from anthropological research into the function of rules, symbolisms, and rituals in various "primitive" societies. Douglas (1966) describes a wide range of "pollution beliefs," held by the society's members, about dangers that may be inflicted on a transgressor, his kin, or others by exhibiting certain purity- or sacredness-defiling behavior. An example is the belief that a women's adultery causes bodily pains in the transgressor's husband. Because this particular society's layered structure is regulated by considerations like marriage payments and marital status, and adultery threatens such established regulatory mechanisms, the argument of Douglas is that such pollution beliefs may very well serve the

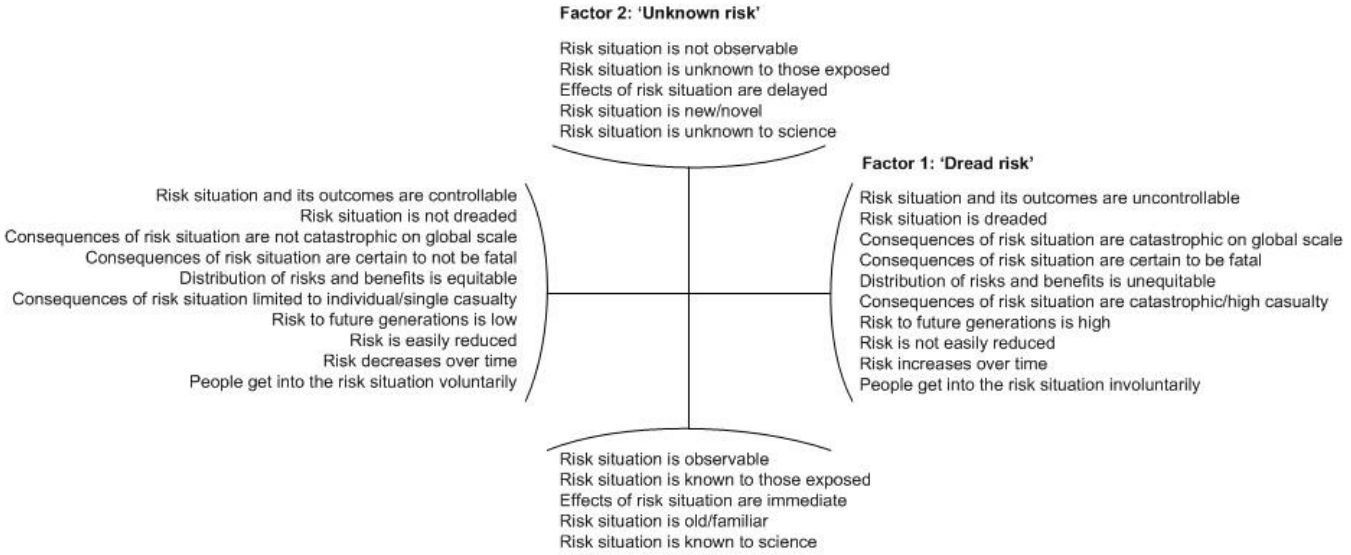

Figure 2.4: Overview of characteristics of risk that make up two key factors in the psychometric paradigm: dread risk and unknown risk. The third factor, number of people exposed, is not shown here.

function of upholding and maintaining the existing social organization of the society. In later work, insights into the relationship between cultural beliefs and the social organization of a group has been applied to "modern societies" to explain differences in people's risk perceptions (see e.g., Douglas & Wildavsky, 1983).

From a systematic analysis of earlier anthropological work, Douglas (1970) discerned two dimensions of "social organization": *group* and *grid*. *Group* describes the degree an individual is incorporated within a bounded social unit, whereas g*rid* describes the degree of social prescriptions and externally imposed rules, such as behavioral rules and established hierarchies among group members. The resulting "grid–group" typology enables conceptual characterizations of the "social organization" of a group because the typology provides four ideal-typical "stable types of transactions between people (Hoppe, 2007, p. 290)": hierarchal groups (high group, high grid), individualist groups (low group, low grid), egalitarian groups (high group, low grid), and fatalist groups (low group, high grid) (see also Figure 2.5).

Besides the anthropological-based grid–group typology, Schwartz and Thompson (1990) argue that a complementary typology can be derived from insights from the field of ecology. This typology, consisting of four "myths of nature," addresses four different perspectives of "ecosystems stability." Specifically, Dake (1992) defines a myth of nature as "one set of beliefs about what the world is like, what its risks are like, and who is to blame for untoward events" (p. 24). The four myths of nature are "nature is capricious," "nature is benign," "nature is tolerant, but within limits," and "nature is ephemeral or fragile" (see Figure 2.6). The key idea is that all four perspectives of ecological stability are technically legitimate, but each proponent of a particular myth of nature will typically view the other myths of nature as irrational (Schwartz & Thompson, 1990). However, some authors problematize the actual degree of correspondence of each of the four myths of nature with one of the four ways of life (see e.g., Grendstad & Selle, 2000).

From the point of view of CTR, regulatory differences can occur due to differences in the cultural beliefs held by influential actors, resulting from differences in these actors' social environment and the organization thereof. The grid–group and myths of nature typologies can function as (separate) concrete yardsticks to discriminate among actors' beliefs about human and physical nature, respectively. For example, we would expect that actors holding individualist beliefs would advocate policies that appeal to these beliefs. Because investments in automobile infrastructure would enhance personal freedom of mobility, we would expect individualists to support such risk management measures. By contrast, the associated increased vehicle capacity of roads would benefit the car users themselves, but could negatively affect nearby residents through increased sound and air pollution. Because this would involve unequal distributions of risks and benefits, thereby undermining egalitarian beliefs, we would expect egalitarian groups to oppose such management measures.

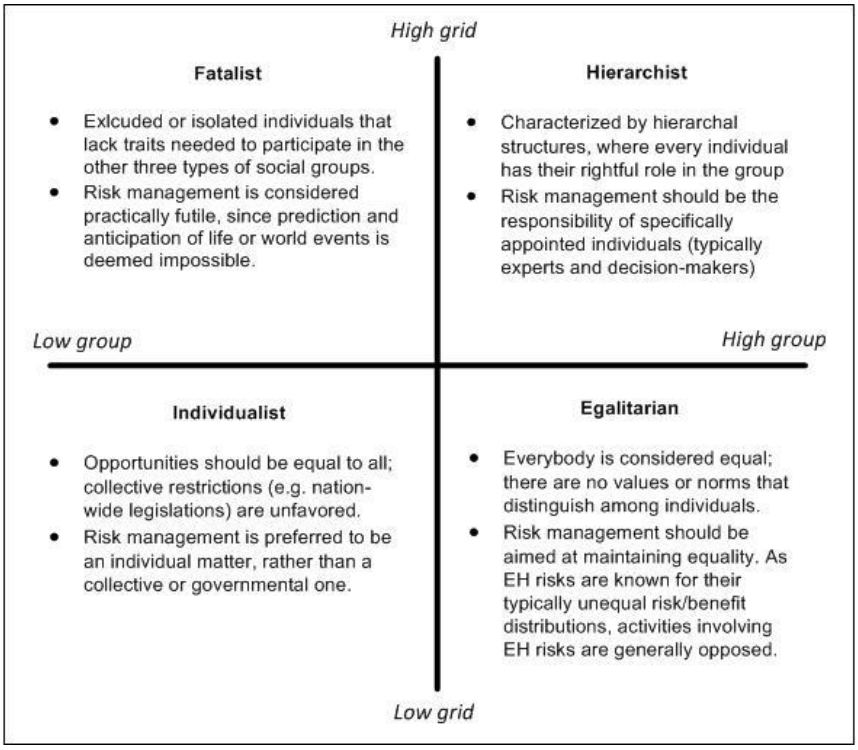

Figure 2.5: Cultural Theory of Risk's grid-group typology, using the 'grid' and 'group' concepts as the y- and x-axis, respectively.

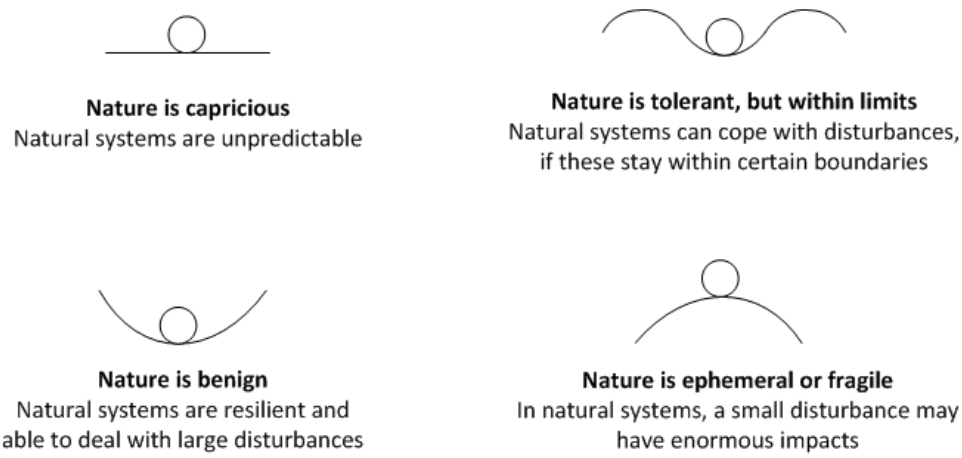

Figure 2.6: Cultural Theory of Risk's Myth of Nature typology, represented by a landscape (the type of natural or human physiological system) and a ball (behavior associated with the risk).

2.3.2.5 Participatory Approaches to Risk Assessment and Risk Management

The RAP has proven an indispensable tool to support evidence-based decision making. However, a string of controversies over the assessment of high-profile chemical and physical risks has unveiled some of its vulnerabilities (see e.g., Jasanoff, 1990). From the 1990s onward, the U.S.-based National Research Council published several reports on how risk assessments should remain credible and authoritative in times of scientific uncertainty and strong competing interests. In 2005, the "European" International Risk Governance Council published its white paper, which addressed similar challenges. In the report "Understanding Risk" of the NRC (1996), the main recommendation was to fundamentally reconsider the concept of "risk characterization." Instead of viewing risk characterization as a summary or translation of mainly biomedical risk knowledge, a much broader conceptualization of risk characterization was envisioned. In particular, risk characterizations should be decision-driven, rather than predominantly science-driven, activities. The multifaceted nature of risk should be acknowledged, extending the scope of risk characterization to include risks relevant to interested and affected parties, such as risks to economic well-being and the potential to undermine personal, social, cultural, and ethical values. Subsequently, the problem formulations, concerns, needs, and interests of interested and affected parties should continuously be taken into account and merged into an "analytic–deliberative" process of risk assessment and risk management (see Figure 2.7). Analytic, since decisions should be informed by rigorous, state-of-the-art science, and deliberative, since interested and affected parties should participate in all phases of the risk assessment process.

Subsequently, in the report "Science and Decisions" of the NRC (2009), the National Research Council revisited the position of risk assessment in contemporary decision making. This volume mainly discusses how the process of risk analysis can be shaped in such a way that risk assessment outcomes provide maximum utility to risk management officials. This time, the 1983

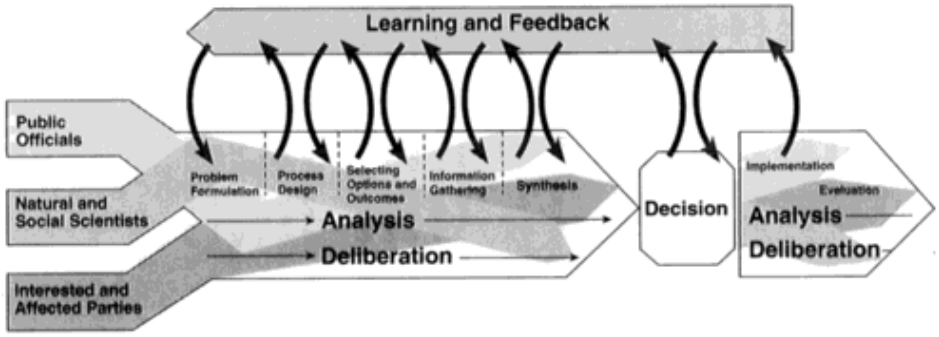

Figure 2.7: The analytic-deliberative process as proposed in the NRC's 1996 'Understanding Risk' report. Source: Ref. NRC, 1996. Permission to reprint this figure has been obtained from the copyright holder (U.S. National Academies Press)

version of the RAP reemerged, but it was explicitly placed in the broader context of political and societal decision-making processes. The NRC recommends to collectively develop problem definitions and an overview of reasonably foreseeably risk management options before risk assessments are formally planned and conducted. Accordingly, actors should be involved before, during, and after the risk assessment by using systematic participatory processes, similar to the recommendations of the NRC's 1996 report. For an overview of this detailed framework, please see NRC (2009).

In Europe, similar calls for the participation of interested and affected parties in risk management were made by both governmental (SCHER/SCENIHR/SCCS, 2013) and independent scientific bodies (IRGC, 2005). Notably, the International Risk Governance Council (IRGC) developed a procedural framework that outlines an inclusive approach to the governance of risks. The framework includes a "risk handling chain" consisting of four phases: preassessment, risk appraisal, tolerability & acceptability judgment, and risk management (IRGC, 2005). The IRGC framework (see also Figure 2.8) views analytical risk research and risk perception research as complementary. Among others, this follows from the "risk appraisal" phase, which includes both a technical risk assessment and a concern assessment. In addition, responsible decision making on systemic risks requires appropriately designed and systematic actor participation, in accordance with the degree of complexity, uncertainty, and ambiguity that characterizes the risk-related knowledge (Hermans, Fox, & van Asselt, 2012; IRGC, 2005, 2008; Renn, 2008; van Asselt & Renn, 2011). According to Renn (2008), "complexity" should here be understood as the difficulty to identify and quantify causal relationships between a variety of potential hazards and the multitude of potential effects following exposure. "Uncertainty" pertains to a situation where the type or nature of any adverse effects, or the likelihood of these effects, cannot be described precisely. Finally, ambiguity refers to a situation where several legitimate and meaningful interpretations of accepted risk assessment results coexist. Nanotechnology risks (IRGC, 2006) and risks posed by synthetic biology (IRGC, 2009) are examples of risks that have been analyzed using a risk governance perspective.

In sum, risk management in democratic societies requires the involvement of interested and affected parties in the various steps of risk assessment and risk management, and explicit and continuous attention to their risk perceptions, including concerns, interests, and needs. The analytical approach to risk is still considered indispensable, though it is seen as one of the necessary ingredients for sound risk management. From the point of view of these frameworks, regulatory differences could occur due to differences in the degree of inclusion of interested and affected parties in risk management and the range of risk perceptions, concerns, interests, and needs voiced by these parties included within risk management.

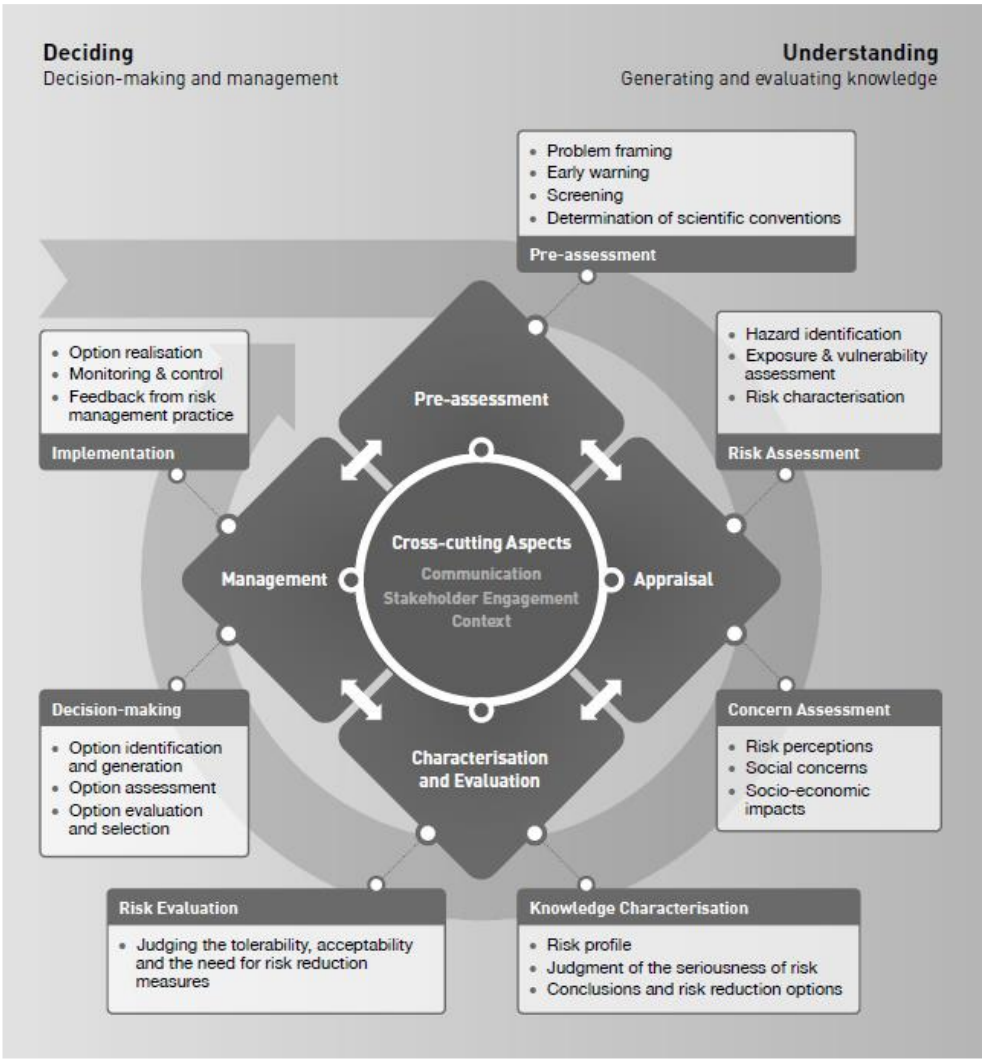

Figure 2.8: The Risk Governance Framework as proposed by the IRGC. Source: IRGC, 2017. Permission to reprint this figure is automatically granted by the copyright holder (EPFL International Risk Governance Center) under condition of acknowledgment.

2.3.2.6 Advocacy Coalition Framework

According to the founders of the Advocacy Coalition Framework (ACF), Sabatier and Jenkins-Smith (1993), the ACF provides a theoretical lens to investigate complex public policy questions. The principal aim of the ACF is to simplify the complexity of public policy for research purposes because public policy issues typically include many different actors with different beliefs and interests, uneven power relations between these actors, and uncertain scientific knowledge (Weible, Sabatier, & McQueen, 2009).

The ACF itself is a comprehensive framework (see also Figure 2.9), with a wide variety of proposed causal relationships, several testable hypotheses, and a set of underlying assumptions. For reasons of brevity, we will focus on two essential concepts of the ACF: the *policy belief system* and the *advocacy coalition*.

A policy belief system is a three-tiered, hierarchal system of beliefs held by those actors that deal with the policy issue on a professional basis (i.e., "policy elites") (Sabatier & Jenkins-Smith, 1993). The policy belief system is a key element of the ACF because the ACF explicitly identifies the beliefs held by politically active actors as the driving force for their political behavior (Weible et al., 2009). A policy elite's policy belief system consists of broad and fundamental deep core beliefs, issue-specific policy core beliefs, and highly specific, more instrumental secondary beliefs (see Figure 2.10).

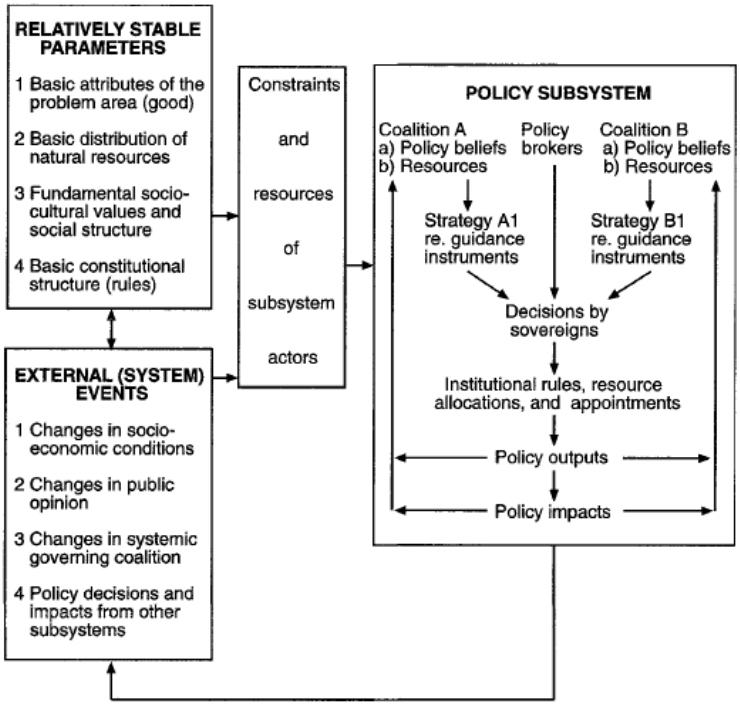

Figure 2.9: The Advocacy Coalition Framework. Source: Sabatier, 1998. Permission to reprint this figure has been granted by the copyright holder (Taylor & Francis).

Though each individual actor is assumed to possess a policy belief system, actors often share certain beliefs. The ACF hypothesizes that when actors share policy core beliefs, they will be inclined to form an advocacy coalition. Here, the term "coalition" should be understood as a circumstantial, loosely connected group of stakeholders that cooperate to a certain degree to have their preferred policy objective implemented. In practice, various advocacy coalitions with competing policy preferences will coexist.

Originally, the ACF was put forward as an alternative to a dominant earlier theory in (public) policy research, the Policy Cycle (also known as the "stages heuristic"; see e.g., Jann &Wegrich, 2007). The ACF may be considered part of the (public) policy studies tradition, whereas the framework also draws, on differing levels, from other disciplines. Prime examples are several theories and broader insights derived from psychology, such as the theory of reasoned action and cognitive dissonance theory (Sabatier, 1998; Sabatier & Jenkins-Smith, 1993). The ACF has been used to analyze various national and international policies, such as financial/economic policies, social policies and education policies, and (public) health and environmental policies, such as air pollution policies and tobacco/smoking policies (Jenkins-Smith, Nohrstedt, Weible, & Sabatier, 2014; Sabatier, 1998; Sabatier & Jenkins-Smith, 1993; Weible & Sabatier, 2007).

From the point of view of the ACF, regulatory differences could occur through the prominence of different advocacy coalitions within different countries for the same policy area. Identifying the deep core beliefs, policy core beliefs, and relevant secondary beliefs of the key actors will then be instrumental in discerning these advocacy coalitions.

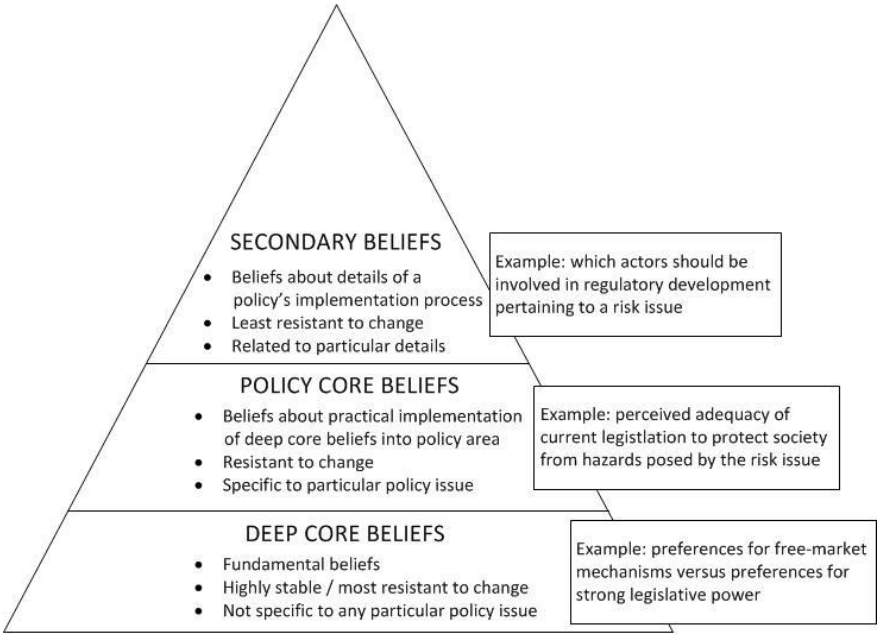

Figure 2.10: The three layers of beliefs, including examples, which together form a hierarchal policy belief system.

2.3.2.7 Social Amplification of Risk Framework

The concept of social amplification of risk was outlined by Kasperson and colleagues in their Social Amplification of Risk Framework (SARF). SARF sets out to explain why a risk associated to relatively limited hazards, exposure, or adverse effects may evoke strong societal reactions, or vice versa, why risks with a substantial attributed impact (from a technical risk assessment perspective) may be attenuated in society (Kasperson et al., 1988; Kasperson, Kasperson, Pidgeon, & Slovic, 2003). SARF introduces the phenomenon of "social amplification of risk" to explain how risk information is continuously framed and reframed by various societal actors, leading to an increased (amplified) or decreased (attenuated) societal response to the risk (see Figure 2.11). For example, Frewer, Miles, and Marsh (2002) noted that increases of media attention to the risks of genetically modified (GM) foods, sparked by worrisome though unpublished research findings, resulted in increased perceptions of risk and decreased perceptions of benefits of GM foods. When the amount of media reporting on GM foods decreased through time, the occurred amplification similarly appeared to diminish, though perception of benefits of GM foods remained depressed for at least a year after the initial media reporting.

In terms of its disciplinary orientation, SARF is explicitly described as an interdisciplinary framework that aims to integrate several strands of previously isolated risk-oriented research (Kasperson et al., 1988). Accordingly, a key feature is the integration of the technical perspective of risk with psychological, sociological, and cultural perspectives of risk. In turn, the framework tries to reconcile the psychological understanding of risk that is central to the Psychometric Paradigm with the sociological and anthropological understanding of risk developed within the Cultural Theory of Risk (Kasperson et al., 1988; Tansey & O'Riordan, 1999). However, some scholars have problematized the attempt to reconcile these two perspectives of risk (see, e.g., Rayner, 1992).

The SARF may be relevant to environmental health risk management because SARF proposes that risk managers or other decisionmakers may react very strongly, or suddenly change their management strategy, as a result of amplified societal reactions. For example, in the Brent Spar controversy, Shell changed its disposal strategy from deep sea disposal to the allegedly less favorable on-shore dismantling after intense societal uproar, mobilized by Greenpeace (Bakir, 2005; Löfstedt & Renn, 1997). The phenomenon of excessive political reactions in response to amplification processes is similar to the notion of "risk regulation reflex," a topic of research in Dutch political science literature. Stringent, sometimes highly invasive measures may be implemented that go substantially beyond the direct effects involved, due to or in anticipation of public responses following the risk event (WRR, 2012). Then, differences in the occurrence of social amplification and the "risk regulation reflex" among countries may explain international differences in risk management. In country A, a certain risk frame may have resonated sufficiently to produce social amplification, whereas in country B, the risk may have passed by relatively unnoticed (i.e., no or little "risk selection"). In country A, subsequent public pressures to act may cause policymakers to revise current risk management or act in accordance with the "risk regulation reflex," whereas such societal and political dynamics will have eluded country B.

FIGURE 1
SOCIAL AMPLIFICATION AND ATTENUATION OF RISK

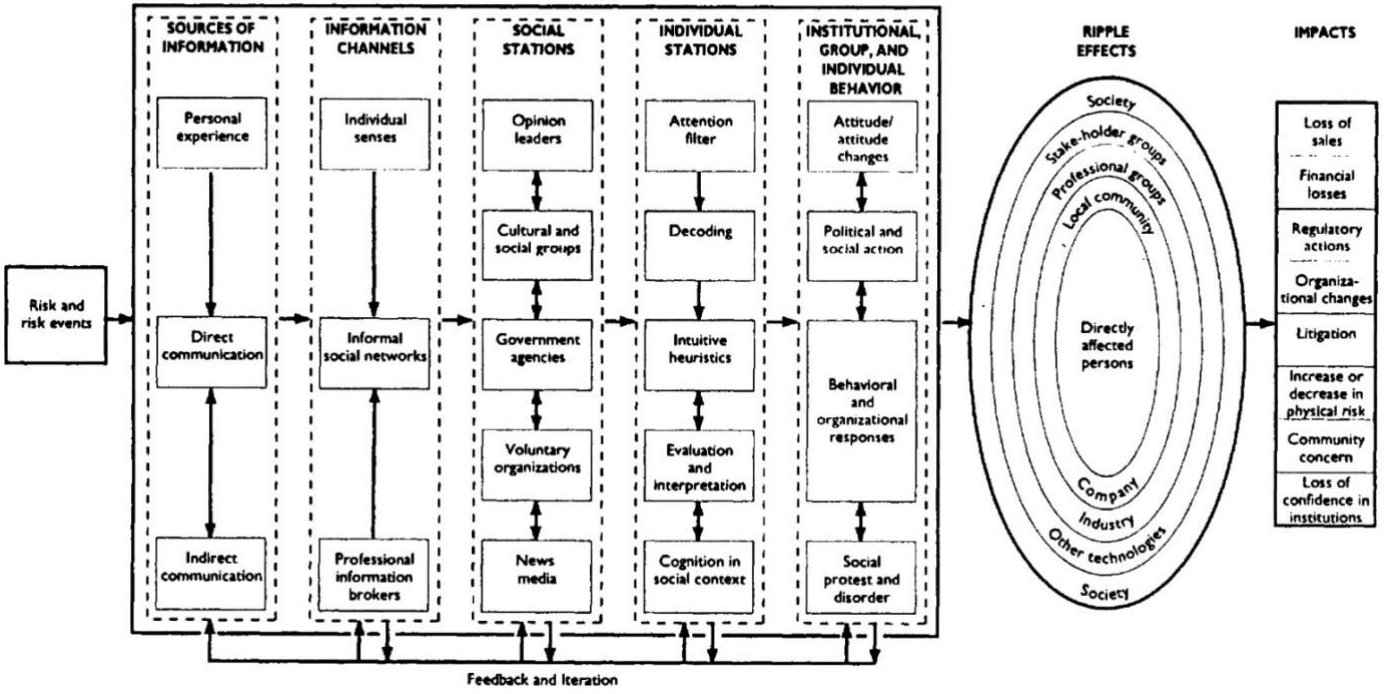

Figure 2.11: The Social Amplification of Risk Framework. Source: Kasperson et al., 1988. Permission to reprint this figure has been granted by the copyright holder (John Wiley and Sons).

2.3.2.8 Hofstede's Model of National Cultures

Hofstede's Model of National Cultures (HMNC) mainly includes six dimensions of national culture, mostly developed from studies into the corporate cultures of subsidiaries of a multinational high-technology company referred to as HERMES (Hofstede, 1980). In the period between 1967 and 1973, Hofstede set out multiple attitude surveys in HERMES subsidiaries in a wide variety of countries (maximum: 67 countries). In general, these surveys included four types of questions, to be answered by the subsidiaries' employees: questions pertaining to employee *satisfaction* (e.g., satisfaction with a certain job aspect), questions asking for employee *perceptions* (e.g., perceived level of stress), questions pertaining to *personal goals and beliefs* (e.g., preferred type of manager), and *demographics* (e.g., years of education). Subsequently, a subset of these questions proved useful in distinguishing four dimensions of national culture: power distance, uncertainty avoidance, individualism, and masculinity. Two dimensions were added in a later stage: long-term orientation (Hofstede, 1991) and indulgence (Hofstede, Hofstede, & Minkov, 2010) (see also Table 2.1).

From the point of view of HMNC, regulatory differences can occur due to differences in national culture, primarily through differences in the characterization of countries' cultures using Hofstede's six dimensions. For example, a more "masculine" country would value entrepreneurship, competitiveness, and financial prosperity and would subsequently emphasize the opportunities offered by potentially risky technological development. We would expect that, in such cultures, general hesitance occurs toward innovation-inhibiting precautionary interventions. Alternatively, a more "feminine" country would emphasize the well-being of the collective, and particularly that of vulnerable groups. In order to inherently protect all populations from potentially hazardous exposures following from a risky technological development, we would expect that such cultures generally prefer (early) precautionary action over "laissez faire" policies.

2.3.3 Key Ideas and Questions Drawn from the Conceptual Frameworks

The eight conceptual frameworks discussed in this article draw attention to different aspects of risk management processes. The second column of Table 2.2 provides an overview of the key ideas introduced by the various frameworks. From these ideas, we derived questions that could guide further empirical work in the form of interviews and document analyses (see the third column of Table 2.2). Whether these questions will actually help understand why certain risk management strategies have been selected, and why such strategies differ between countries, will follow from future empirical research.

Table 2.1: Overview of the six dimensions of national cultures developed by Hofstede and colleagues, including related core values and definitions.

Cultural dimension	**Related core value (after Ref. 69)**	**Definition**
Power distance (high vs. low)	How a society's members find solutions to the fundamental issue of human inequality	… "the extent to which the less powerful members of institutions and organizations within a country expect and accept that power is distributed unequally (p.46)" (Ref. 70)
Uncertainty avoidance (high vs. low)	How stressful the idea of an unknown future is to a society's members	… "the extent to which the members of a culture feel threatened by ambiguous or unknown situations (p.167)" (Ref. 70)
Individualism (vs. collectivism)	How a society's individual members integrate into certain groups	Individualism: a society "in which the ties between individuals are loose: every is expected to look after himself (…) and his (…) immediate family (p.76)." (Ref. 70) Collectivism: a society "in which people from birth onward are integrated into strong, cohesive in-groups, which throughout people's lifetimes continue to protect them in exchange for unquestioning loyalty (p.76)" (Ref. 70)
Masculinity (vs. femininity)	How the division of emotional roles of men and women has taken shape in a society	Masculine society: … "emotional gender roles are clearly distinct: men are supposed to be assertive, tough, and focused on material success, whereas women are supposed to be more modest, tender and concerned with the quality of life (p.120)" (Ref. 70) Feminine society: … "emotional gender roles overlap: both men and women are supposed to be modest, tender, and concerned with the quality of life (p.120)" (Ref. 70)
Long Term Orientation (vs. short term orientation)	How a society's members direct their efforts timewise	Long-term orientation: "fostering of virtues oriented toward future rewards; in particular, perseverance and thrift (p.210)" (Ref. 70) Short-term orientation: "fostering of virtues related to the past and present; in particular respect for tradition, preservation of `face` and fulfilling social obligations (p.210)" (Ref. 70)
Indulgence (vs. restraint)	How a society's members go about satisfying their wants and needs to enjoy life	Indulgent society: "allows relatively free gratification of basic and natural human desires related to enjoying life and having fun (p.15)" (Ref. 69) Restrained society: "controls gratification of needs and regulates it by means of strict social norms (p.15)" (Ref. 69)

Table 2.2: The eight conceptual frameworks and their properties to explain variation in international risk management strategies, including questions derived from these frameworks that could be used to empirically study these variations. The questions pertaining to Hofstede's model of national culture are presented in italic, since these questions cannot be used directly during interviews.

Name and (disciplinary) orientation of conceptual frameworks	**Key ideas brought forward by the frameworks**	**Questions relevant to the analysis of international differences in environmental health risk management**
Risk Assessment Paradigm (RAP); analytical perspective used by subject-matter experts	> Preferably quantitative assessment of risk numbers, based on toxicological working mechanisms, dose-response knowledge and exposure in the population > Separation of risk assessment (life science oriented) and risk management (normative, policy oriented balancing of various different perspectives) > Often interpretative ambiguity about hazard- and risk science base in risk assessment	> How was risk assessment knowledge weighted against other perspectives? > Was uncertainty addressed, and how did uncertainties affect risk management? > Were safety or assessment factors used to define safe levels? > Were availability and risks of alternatives considered? > Was precaution addressed explicitly?
Advisory roles of experts in risk management processes (ARERMP); interdisciplinary field of research dealing with roles of subject-matter experts in risk management processes	> Experts can take different roles when providing advice to policy makers. Ideal typical expert roles include pure scientist, science arbiter, issue advocate and honest broker of policy alternatives > Personal views of experts affect their willingness to advise on different action perspectives. Ideal-typical expert roles include science absolutist, technological optimist, environmental centrist, cautious environmentalist, environmental absolutist > Expert roles are strongly context dependent, depending on e.g. the characteristics of the issue and the organization the expert works in	> Which expert roles can be identified in the risk management process? > Were specific expert roles dominant? > How could the occurrence of these roles have affected risk management process? > Would inclusion of other expert roles have allow for other action perspectives? > To what extent were different expert roles present in the expert commission or decision-making team?
Psychometric Paradigm (PP); psychological perspective on risk perception of the general public	> Evaluations of risks depend on how individuals construct risks > Focus on acceptability of risk by individuals, not on risk numbers applying uniformly to any individual > Several dimensions of risk acceptability defined, e.g. involuntary nature of risk, equitable distribution of risks and benefits, perceived controllability > Actors use heuristics in the evaluation of information and of risks	> How does the general public perceive the risk? > Could these perceptions be characterized as 'dread risk' or 'unknown risk'? > How may these risk perceptions have impacted risk management?

Table 2.2: Continued

Cultural Theory of Risk (CTR); sociological and anthropological perspective on risk perception of lay people and culturally determined (core) beliefs	> Characterization of two dimensions of social organization, i.e. grid and group; four ways of life follow from different types of social organization: egalitarian, individualist, fatalist, hierarchist. These four views characterize attitudes to acceptability of risks and preferred management of risks > Four different perspectives on 'ecological stability', or myths of nature: nature is benign, nature is ephemeral or fragile, nature is robust, but within limits, nature is capricious. These four beliefs characterize attitudes to acceptability of risks and preferred management or risk > A relationship between the four ways of life and the four myths of nature is hypothesized	> Do different groups hold different perceptions of risk? > To what extent did arguments in the risk management process resemble any way of life? > To what extent did arguments in the risk management process resemble any myth of nature? Were one of these perspectives dominant in the decision-making discourse? > How could this have affected the decisions made? > Did the composition of the expert commission or decision-making team sufficiently cover the different ways of life and myths of nature perspectives?
(Hybrid) participatory approaches to risk assessment and management (PARAM); Procedural frameworks combining the technocratic analytical approach to risk assessment with insight from 'social scientific' risk perception research	> Stakeholder involvement in risk assessment, management and governance > Concern assessment parallel to risk assessment (IRGC framework) > Decision contexts and societal needs should be the starting point for assessments and characterizations of risk	> Which actors have been involved in the risk management process? > To what extent have perceptions of risk been taken forward in the decision-making process? > What specific policy or societal questions were the risk assessments required to address?

Table 2.2: Continued

Advocacy Coalition Framework (ACF); perspective from policy sciences	> Beliefs held by (professional) actors can be structured in a hierarchal, three-layered policy belief system, consisting of deep core beliefs, policy core beliefs, and secondary beliefs > Actors holding similar policy core beliefs will cooperate in advocacy coalitions to have their preferred policy measures implemented > Various advocacy coalitions will typically coexist and attempt to advance their political agenda	> What are the policy belief systems of the most influential actors? > What advocacy coalitions have been present in the decision-making process? > Has a certain advocacy coalition been dominant in the decision-making process? > What were the most important (core) beliefs of the actors cooperating in the dominant advocacy coalition?
Social Amplification of Risk Framework (SARF); integrative perspective drawing from multiple fields of risk research	> Information processing and interactions between different actors determine the extent of social amplification or attenuation of risks > Actors emphasize different parts of messages, leading to different framings and different evaluations of information and of risks	> Did certain risk frames resonate with particular actors, including members of the general public? > Did this resonance induce amplification or attenuation? To what extent may this amplification have affected the decision-making process?
Hofstede's model of national culture (HMNC); perspective focusing on how cultures differ between countries	> The culture of individual countries can be characterized in terms of six dimensions: power distance, uncertainty avoidance, individualism, masculinity, long-term orientation, indulgence	*> How does the country score on the six dimensions of Hofstede?* *> Are these differences between the analyzed countries, with respect to the scoring on certain dimensions?* *> To what extent do potential differences in scoring relate to the observed differences in risk management strategies?*

2.4 DISCUSSION AND CONCLUSIONS

In this article, we have explored a range of theoretical perspectives applicable to the analysis of reasoning behind divergent risk management strategies in the environmental health domain. This exploration was performed while explicitly acknowledging the often complex and intimate relationships between science and policy. Ultimately, our aim is to identify, understand, and compare the arguments used to develop national-level decision making in various countries, for the same risk, and how, and to what extent scientific expertise is involved in shaping these arguments.

2.4.1 Overlap and Differences in Disciplinary Orientations

When comparing the disciplinary orientations of the eight conceptual frameworks (using the first column of Table 2.2), both differences and overlap can be observed. To gain more insight in the areas of science in which the frameworks are used, and how these areas contrast or match with one another, we conducted a literature search in Scopus. For each conceptual framework, we compiled a list of relevant publications using a specific search query (see Table 2.3). We then categorized the resulting list using the 28 subject areas (including "undefined") contained in Scopus and calculated the following ratio for each subject area: the number of publications falling within 1 of the 28 subject areas as compared to the total number of publications found for that specific framework. Table 2.3 shows these ratios, as well as the list of queries used and the 28 subject areas used for categorization. We arbitrarily consider a subject area significant for a particular conceptual framework when at least 5% of the total literature was found to fall into that area of science.

Most notably, we observe that the RAP and HMNC are applied in distinctly different areas of science as compared to the other conceptual frameworks. Table 2.3 shows that the RAP stands out by being the only framework where we found a significant share of literature related to "pharmacology, toxicology, and pharmaceutics," but lacking a significant share of literature related to "social sciences." These findings reinforce the idea that literature referencing the RAP still has a particularly technical and biomedical disciplinary orientation. HMNC stands out by its strong orientation toward "business, management, and accounting" literature, by being the only framework with a significant share of literature in "computer science" and "economics, econometrics and finance," and by missing a significant share of literature for "environmental science" and "medicine." These findings may be unsurprising because HMNC has no substantive links to any type of environmental health-related risk research. Alternatively, the six other frameworks all have significant shares of literature in both the "social sciences" and some of the biomedical-related ("environmental science" and "medicine") areas of science. These findings appear to reinforce the idea that the majority of frameworks included in this article combine social scientific perspectives with case studies related to human or environmental risks. Concluding, this bibliometric comparison does appear to underline the "polythetic" nature of the concept of "risk."

Table 2.3: Upper part: Results from the literature search showing the percentage of publications falling into a particular subject area, as compared to the total number of publications gathered for that conceptual framework; percentages printed in **bold**: assumed significance (percentage ≥ 5%); subject areas printed in **bold**: at least six of the eight conceptual frameworks score ≥5% (assumed significance) for a particular subject area. Lower part: list of the abbreviations used in the upper part of Table 2.3 and the queries used to perform the literature search in Scopus.

	THEORETICAL APPROACHES							
SUBJECT AREAS USED IN SCOPUS	RAP	ER	PP	CToR	IRGC	ACF	SARF	Hofst.
Agricultural and Biological Sciences	**5%**	**5%**	3%	0%	4%	**5%**	1%	0%
Arts and Humanities	0%	**9%**	2%	4%	2%	2%	2%	4%
Biochemistry, Genetics and Molecular Biology	**6%**	3%	1%	0%	1%	0%	1%	0%
Business, Management and Accounting	0%	2%	**7%**	**11%**	**9%**	**6%**	**11%**	**34%**
Chemical Engineering	3%	1%	0%	0%	3%	1%	0%	0%
Chemistry	3%	0%	0%	0%	4%	0%	0%	0%
Computer Science	0%	1%	3%	1%	0%	0%	1%	**11%**
Decision Sciences	0%	0%	1%	2%	1%	0%	1%	4%
Dentistry	0%	0%	0%	0%	0%	0%	0%	0%
Earth and Planetary Sciences	0%	**5%**	3%	3%	**5%**	1%	3%	0%
Economics, Econometrics and Finance	0%	3%	1%	2%	4%	4%	3%	**14%**
Energy	0%	1%	2%	0%	1%	1%	1%	1%
Engineering	**6%**	**5%**	**19%**	**9%**	**19%**	1%	**17%**	4%
Environmental Science	**29%**	**14%**	**12%**	**10%**	**10%**	**24%**	**9%**	1%
Health Professions	0%	1%	0%	1%	0%	0%	0%	0%
Immunology and Microbiology	2%	1%	0%	0%	1%	0%	0%	0%
Materials Science	1%	0%	1%	0%	2%	0%	0%	0%
Mathematics	0%	1%	1%	0%	2%	0%	1%	1%
Medicine	**19%**	**15%**	**15%**	**14%**	**6%**	**5%**	**17%**	1%
Multidisciplinary	0%	1%	0%	0%	0%	0%	0%	0%
Neuroscience	0%	0%	0%	0%	0%	0%	0%	0%
Nursing	0%	3%	1%	1%	0%	1%	0%	0%
Pharmacology, Toxicology and Pharmaceutics	**22%**	1%	1%	0%	1%	0%	0%	0%
Physics and Astronomy	0%	1%	0%	0%	2%	0%	0%	0%
Psychology	0%	1%	**7%**	**7%**	0%	1%	**7%**	4%
Social Sciences	3%	**26%**	**16%**	**31%**	**22%**	**48%**	**23%**	**19%**
Veterinary	0%	1%	0%	0%	1%	0%	0%	0%
Undefined	0%	0%	0%	0%	0%	0%	0%	0%

LEGEND	**QUERY USED IN SCOPUS**
RAP: Risk Assessment Paradigm	TITLE-ABS-KEY("risk assessment paradigm" AND ("hazard" OR "exposure"))
ER: Expert Roles	TITLE-ABS-KEY("Policy" AND "Scien*" AND Expert* AND "Role*" AND "Advi*")
PP: Psychometric Paradigm	TITLE-ABS-KEY("Psychometric Paradigm" AND "Risk")
CToR: Cultural Theory of Risk	TITLE-ABS-KEY("Cultural Theory" AND "Risk")
IRGC: IRGC Risk Governance Framework	TITLE-ABS-KEY("International Risk Governance Council" OR "Risk Governance Framework")
ACF: Advocacy Coalition Framework	TITLE-ABS-KEY("Advocacy Coalition Framework")
SARF: Social Amplification of Risk Framework	TITLE-ABS-KEY("Social Amplification" AND "Risk")
Hofst.: Hofstede's model of national cultures	TITLE-ABS-KEY("Hofstede" AND "National Culture")

Bold percentages: percentage >= 5%

Bold subject area: more than six of the eight approaches score >= 5% for a particular subject area

In terms of limitations, we used one single search engine (i.e., Scopus), we used relatively simple search queries, and we did not exclude any publications from the list of literature obtained using the query. However, we think that our pragmatic approach to a literature search is appropriate to get a taste of the areas of science relevant to each of the eight conceptual frameworks.

2.4.2 Visualization of (Conceptual) Relationships Between the Frameworks

Although there are distinct disciplinary differences between some of the conceptual frameworks, some conceptual overlap can be found between many of them. For example, the four worldviews in the grid/group typology of the CTR could provide a theory-based specification of the fundamental "deep core beliefs" in the ACF (Hoppe, 2007; Jenkins-Smith, Silva, Gupta, & Ripberger, 2014). Our attempt to visualize as many connections as possible between the key concepts of the various conceptual frameworks resulted in one overarching framework (see Figure 2.12).

In short, we propose two levels of analysis: a level of interaction occurring *between* advocacy coalitions, and a level of argumentative and cognitive processes taking place *within* the stakeholder or coalition. The upper level (upper part of Figure 2.12) is mostly inspired by the SARF and the ACF and consists of a timeline, in which the competing advocacy coalitions make their argumentative moves to influence the political course of action. The advocacy coalition whose arguments (and framings contained therein) amplify most strongly in society will be expected to dictate this course of action. Next, we believe that an advocacy coalition's argumentative moves are performed in accordance with predetermined sets of beliefs (bottom part of Figure 2.12). We draw from the ACF to make a distinction between *surface* beliefs that are specific to the risk, and *fundamental* beliefs that transcend specific issues. We identify specific types of fundamental beliefs: the five expert roles developed by Weiss (2003) and the four myths of nature and four cultural worldviews developed within the framework of the CTR. Next, we distinguish scientific/professional reasoning and knowledge from normative reasoning and personally held beliefs. We propose that professional issue-specific knowledge and beliefs are influenced by the various determinants of expert roles outlined by Spruijt et al. (2014). By contrast, personal knowledge and beliefs are influenced by social norms, community culture (cf. CTR), and the characteristics of one's national culture (cf. HMNC).

In terms of limitations, some connections between frameworks may have been missed or may have been overinterpreted, and this framework is not supported by empirical evidence. However, the framework does provide, in one relatively simple figure, a visualization of the various concepts and ideas that may be relevant for the study of divergent risk management strategies in the environmental health domain.

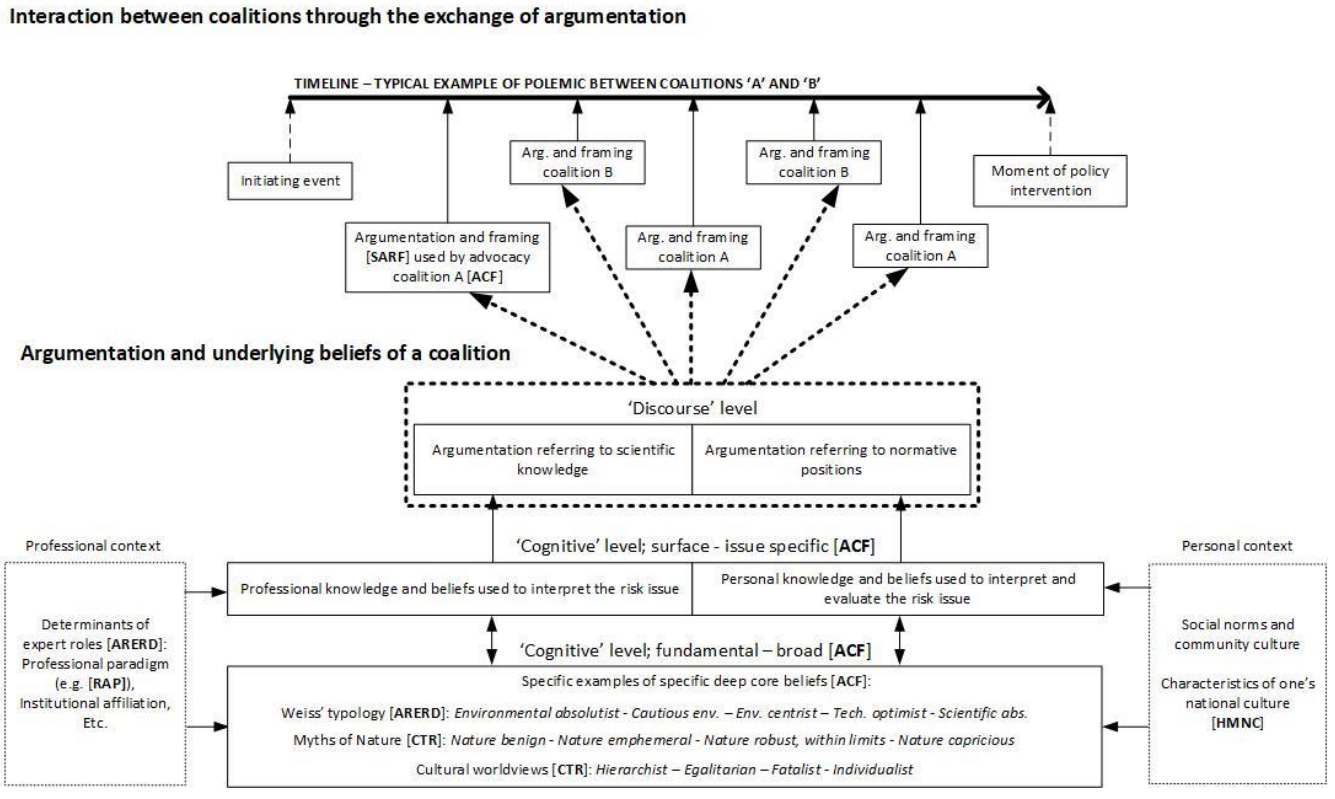

Figure 2.12. An overarching framework that visualizes the various (conceptual) relationships between the eight conceptual frameworks. Abbreviations of the names of the frameworks, printed in [**BOLD CAPITALS**] and in between square brackets, are used to signal that an idea or concept of that respective framework has been used as a source of inspiration for that specific section of the overarching framework. List of the abbreviations used: [**SARF**], Social Amplification of Risk Framework; [**ACF**], Advocacy Coalition Framework; [**ARERMP**], Advisory Roles of Experts in Risk Management Processes; [**RAP**], Risk Assessment Paradigm; [**CTR**], Cultural Theory of Risk; [**HMNC**], Hofstede's Model of National Cultures.

2.4.3. Strengths and Limitations

As far as we are aware, this article is the first study where conceptual frameworks from a wide variety of academic disciplines are applied to seek explanations for differences in environmental health risk management strategies between countries. We argue that these conceptual frameworks are complementary to one another, by offering different concepts and ideas to study these differences. Indeed, Slovic (1999) wrote that "whoever controls the definition of risk controls the rational solution to the problem at hand" (p. 689), thereby drawing an explicit link between the numerous perspectives of risk and how these impact what is considered to be, and how one should arrive at, "rational" risk management strategies.

In this article, we are mainly interested in differences in environmental health risk management between countries. However, within countries, risk management strategies can also differ from policy area to policy area (see e.g., Hood, Rothstein, & Baldwin, 2001). For example, Johannesson, Hansson, Rudén, and Wingborg (1999) showed that in Sweden, major differences exist between the policy areas of occupational safety and health, environmental protection, and chemicals control, in terms of the organization of legal mandates across the relevant agencies, the regulatory strategies followed and the enforcement activities involved. Although we acknowledge the typical diversity of risk management strategies within a country, we argue that discussions about how and why such *intranational* diversity occurs are beyond the scope of this article.

In addition, we realize that the conceptual frameworks discussed in this article offer inherently simplified representations of the complexity and dynamics associated with risk management processes. A key contextual factor that is not explicitly part of this theoretical review is how preconditions and regulatory boundaries set by existing regulatory frameworks could influence risk management processes. Regulatory frameworks in the environmental health domain typically provide standards (at least partially) based on scientific evidence, and provisions for the generation of specific types of scientific data (minimally) necessary to evaluate whether the potential hazard meets the established safety standards. These arrangements are developed based on negotiations between the interested and affected parties. These parties can strategically use the room offered by the resulting regulatory arrangements to serve their specific interests. For example, various parties prefer to reduce the amount of animal tests (for various reasons), and each party can use the provisions available in a regulatory framework to cater to this preference. A further discussion of such stakeholder maneuvering in the context of particular regulatory frameworks is outside the scope of this review of conceptual frameworks. However, future case studies should take into account the regulatory context, and particularly the associated uses of scientific evidence, in which the risk management process is taking place.

2.5 CONCLUSIONS

We have analyzed a variety of conceptual frameworks in order to develop a list of questions we can use to study divergent (preferred) risk management strategies in the domain of environmental health risks. We were specifically attentive to the role of science in risk management processes. In future research, the analytical value of these questions will be tested empirically through case studies in the field of environmental health risk management.

ACKNOWLEDGMENTS

The authors would like to thank Jeroen Devilee and two anonymous reviewers for their insightful comments on an earlier version of this article. The project was financed in the framework of the RIVM Strategic Programme (S/124002).

Understanding conflicting views of endocrine disruptor experts: a pilot study using argumentation analysis

Sander C.S. Clahsen, Holly S. van Klaveren, Theo G. Vermeire, Irene van Kamp, Bart Garssen, Aldert H. Piersma and Erik Lebret

Published in:
Journal of Risk Research
Volume 23, Issue 1, 2020

ABSTRACT *To what extent do substances have the potential to cause adverse health effects through an endocrine mode of action? This question elicited intense debates between endocrine disrupting substances (EDS) experts. The pervasive nature of the underlying differences of opinion justifies a systematic analysis of the argumentation put forward by the experts involved. Two scientific publications pertaining to EDS science were analyzed using pragma-dialectical argumentation theory (PDAT). PDAT's methodology allowed us to perform a maximally impartial and systematic analysis. Using PDAT, the structure of the argumentation put forward in both publications was reconstructed, main standpoints, and arguments were identified, underlying unexpressed premises were made explicit and major differences in starting points were uncovered. The five differences in starting points identified were subdivided into two categories: interpretative ambiguity about underlying scientific evidence and normative ambiguity about differences in broader norms and values. Accordingly, two differences in starting points were explored further using existing risk and expert role typologies. We emphasize that particularly the settlement of normative ambiguity, through the involvement of broader ethical, social or political values, inherently requires multi-stakeholder approaches. Extrapolation of our findings to the broader discussion on EDS science and further exploration of the roles of EDS experts in policy processes should follow from further research.*

3.1 INTRODUCTION

The toxicological and epidemiological research of hazards and risks has proven indispensable in supporting evidence-based decision-making in the environmental health domain. A wide range of regulatory agencies and scientific institutes rely on the principles of this research to derive safe limits for the exposure to all kinds of chemical and physical agents. However, this does not mean that experts always agree about the interpretation and evaluation of the available evidence.

There are numerous examples of environmental health risk issues where experts disagree about whether apparent exposure levels can in fact adversely impact public health or the environment. For example, it remains debated whether the available evidence shows that current exposure levels to electromagnetic fields may cause detrimental effects in humans (compare e.g. Sage, Carpenter, and Hardell 2016 and SCENIHR 2015). Expert disagreement about the carcinogenic properties of glyphosate (compare EFSA 2015 and FAO/WHO 2016, and IARC 2017) has made the herbicide's regulatory (re)approval in the EU quite controversial (Science 2016). These are just a few examples where experts differ in their interpretation of scientific evidence surrounding a risk issue.

In this article, we are interested in the differences of opinion occurring between experts in the scientific debate on endocrine disrupting substances (EDS). Beronius et al. (2010) discussed the impact of risk assessment methodologies and expert judgment on the process and outcomes of risk assessments of Bisphenol A. However, the values held by experts that could potentially influence their scientific judgments were not further evaluated. The influence of values and worldviews on the substance of experts' policy advice and the dynamics of scientific discussions in general have been studied extensively (see e.g. Douglas 2000; Elliott 2017; Pielke, 2007; Sarewitz 2004; Spruijt et al. 2014), though we are unaware of studies that specifically identify differences in values at stake in the EDS controversy. Therefore, we aim to analyze some of the values that appear to play a role in the various expert positions in the scientific debate on EDS risk and to identify whether these are based on different interpretations of the underlying scientific knowledge, or whether normative (value) judgments are involved. In the policy and (risk) governance sciences, the distinction between (un)certainty in science and knowledge on the one hand and (lack of) consensus on norms and values on the other hand is important, since these are thought to require different strategies to conflict settlement. The credo "more research is needed" applies to uncertainty in knowledge, while dissensus on norms and values requires multi-stakeholder approaches (Hisschemöller and Hoppe 1995; Renn 2008). Examples of such approaches are participatory discourses (IRGC 2005; Renn 2008) or 'extended peer-community' approaches (see e.g. Ravetz 1999).

Drawing from the existing distinction between uncertainty in knowledge and dissensus on norms and values, we use a classification provided by Renn (2008) to distinguish between two types of ambiguity, interpretative and normative ambiguity. Interpretative ambiguity refers to different interpretations of specific pieces of scientific evidence, for example about the meaning of toxicological evidence versus epidemiological evidence, or weight-of-evidence. Normative

ambiguity refers to differences in values and ethical norms, for example about the acceptability or tolerability of a risk. Note that these two types of ambiguity are not mutually exclusive, and that a certain difference of opinion can contain elements of both (see e.g. Renn 2008). Interpretative ambiguity could, for example, involve normative judgments about the adversity of effects (i.e., whether an observed outcome from a toxicological study constitutes an adverse effect or not). Subsequently, one could argue that, in such cases, interpretative ambiguity would also require the attention of those affected by the risk issue. By contrast, due to the nature of the (often ethical) values involved, we argue that multi-stakeholder approaches are inherently required to settle normative ambiguity, or conversely, that it would be ineffective to settle this type of ambiguity solely in the scientific sphere.

To identify interpretative and normative ambiguity in the scientific debate on EDS risk, we use the pragma-dialectical argumentation theory (van Eemeren, Grootendorst, and Henkemans 2010; van Eemeren and Grootendorst 1984) as the framework to analyze two scientific publications from the realm of regulatory science (i.e., Lamb et al. 2014; Bergman et al. 2015). These two publications touch upon a wide variety of topics relevant for EDS science. We demonstrate the added value of the systematic analysis of argumentation to explore differences in the values expressed by experts. These differences may subsequently clarify why some differences of opinion among EDS experts appear to be pervasive.

Two questions related to the scope of our research may emerge: (1) Why focus on the debate of EDS science? (2) Why use argumentation theory as the analytical 'lens' to study this debate? First, the potential of substances to cause adverse health effects through an endocrine mode of action remains an intensely debated area (see e.g. Autrup et al. 2015; Zoeller et al. 2014). A widely accepted definition of an EDS has been proposed in the 2002 WHO-IPCS state-of-the-science report: 'an endocrine disruptor is an exogenous substance or mixture that alters function(s) of the endocrine system and consequently causes adverse health effects in an intact organism, or its progeny, or (sub)populations' (WHO-IPCS 2002). However, the practical application of this definition is surrounded by controversy. It is not uncommon for experts to disagree about important aspects of EDS science, such as aspects pertaining to the identification of EDS:

- What evidence is specifically required and to what extent will the required scientific evidence about adverse health effects, modes of action and the causal link between the two be available?
- What does evidence in 'intact animals' constitute precisely and how should results from *in vitro* studies or other non-animal studies be considered?
- To what extent do EDS comply with the assumption that adverse physiological effects follow a threshold mechanism?

Experts in the field of EDS science may answer differently to these questions. We think an approach is justified that systematically analyzes the argumentation put forward by the various experts and expert groups that take a position in this debate, for various reasons. First, an argumentation analysis provides insight into the complex argumentation put forward in a (scientific) discussion. This insight is needed to say anything about the structure of the

argumentation put forward and the way the discussion proceeds. Making use of the analytical tools provided by the pragma-dialectical argumentation theory, an analyst can identify the exact standpoint(s) under discussion and any supporting argumentation. The analyst could also describe to what extent parties respond to one another's standpoints and argumentation. Also, implicit elements in the argumentation (standpoints or premises) can be made explicit. Second, an argumentation analysis could be used to assess the quality of a discussion (i.e., whether rules for an argumentative discussion are not violated), though this type of analysis will not be performed in this article. Third, an argumentation analysis may illuminate prerequisite knowledge, professional experiences or personal beliefs (i.e., starting points, see Table 3.1) that remain implicit, but should nevertheless be made explicit to understand the origin of a discussant's standpoint and the supporting argumentation. However, the only way to systematically identify such starting points is to use the aforementioned analytical tools to reconstruct the argumentation structure, and then subject this argumentation structure to thorough analysis. Identified substantive points of departure in the discussion are further distinguished in interpretative ambiguity from normative ambiguity.

Finally, an in-depth argumentation analysis of publications in the realm of regulatory science requires consideration of two prerequisites. First, the credibility of such an analysis will strongly depend on the impartiality of the analysts, so (inherently normative) judgments about the truth value of the premises used in the argumentation put forward in both publications are outside the scope of this article. Second, such a neutral position is particularly relevant when analyzing 'regulatory science' publications. One could evaluate to what extent argumentation put forward is in accordance with 'regulatory reality'. That is, various regulatory frameworks concerned with the safety of industrial chemicals or pesticides have data-requirements to enable science-based risk management. However, a report from the RIVM (2016) has shown that for EDS, the current minimal data requirements in EU's relevant regulatory frameworks are insufficient to identify an EDS based on the currently proposed (science-based) EU criteria. To maintain a neutral position in our argumentation analysis, we refrain from evaluations based on such contextual information and strictly focus on the specific argumentation (and related unexpressed premises) put forward in the two publications. For the present analysis, this means that the focus is limited to a discussion about the scientific merits of the WHO/UNEP 2012 report, as addressed by the argumentation put forward in the two publications selected for analysis in this article. On the basis of the aim and scope of our study, we have articulated four main research questions (1a, 1 b, 1c and 1d) and one sub question (2):

1a. *How does the discussion between Lamb et al. (2014) and Bergman et al. (2015) proceed in argumentative terms?*
1b. *Based on this argumentation analysis, how do the starting points identified from the Lamb et al. (2014) publication and the starting points identified from the Bergman et al. (2015) publication differ from each other?*
1c. *To what extent do these differences in starting points pertain to interpretative or normative ambiguity?*
1d. *Using this classification of the identified starting points, what types of value differences appear to be at stake in the analyzed discussion?*

2. *How does argumentation theory provide additional value through its ability to explicate and clarify obstructions to the various differences of opinion in the field of EDS science?*

3.2 THEORY OF ARGUMENTATION ANALYSIS

To perform an argumentation analysis systematically, we use the pragma-dialectical argumentation theory (PDAT), a theory of argumentation developed by van Eemeren and colleagues (van Eemeren, Grootendorst, and Henkemans 2010; van Eemeren and Grootendorst 1984). PDAT is a well-known approach in the area of argumentation theory and was developed as a response to an experienced lack of a systematic approach to study argumentation in a wide variety of (social) settings, e.g. argumentation in 'every-day life', argumentation in legal settings or argumentation in scientific discussions. PDAT aims to include and build upon concepts from other schools of thought studying argumentation. Since its inception in the beginning of the 1980s, the theory was further developed through discussions among scholars in argumentation theory and beyond, exemplified by the wealth of literature available (see e.g. The Handbook of Argumentation Theory of van Eemeren et al. 2014).

According to PDAT, the goal of an argumentative discussion is to solve a difference of opinion by means of reasonable argumentation. Resolving a difference of opinion means that two discussants jointly see if a standpoint is tenable against criticism. For more detailed information on PDAT we refer to the Supporting Information. Here only the concepts and key steps of analysis are summarized. Note that some of the key concepts of PDAT are named after terms that are also used in 'common language'. Also, some of these concepts may appear to be similar to one another, while the same concept may be understood differently within different schools of thought. Because the concepts used throughout this article bear specific meaning (following PDAT terminology) and have been developed in line with traditions in the discipline of argumentation theory, these concepts should be interpreted accordingly. Table 3.1 provides an overview of key concepts and their specific uses in PDAT. Throughout this article, we use italics when we refer to one of the concepts of PDAT to avoid ambiguity as much as possible.

We specifically focus on two vital steps in the analysis of an exchange of argumentation. The first step of our analysis is identifying the stages of the ideal model of a critical discussion in the actual argumentative contributions of both parties involved (van Eemeren, Grootendorst, and Henkemans 2010). This model consists of four stages that discussants should ideally follow when participating in a critical discussion: the confrontation stage, opening stage, argumentation stage and concluding stage (see Appendix Table A3.1 for more detailed information).

The second step is identifying the standpoints and the argumentation, making explicit of implicit elements in the argumentation, reconstructing the argumentation structure and identifying the argument schemes used. PDAT distinguishes between the external and internal organization of argumentation, referring to 'argumentation structures' and 'argument schemes', respectively (van Eemeren et al. 2014). Argumentation structures provide a complete

Table 3.1: Key concepts and their uses in pragma-dialectial argumentation theory.

Concept	Uses in pragma-dialectical argumentation theory (PDAT)
Standpoint	Refers to what is at issue in argumentative discourse, i.e., what is argued about by the parties (van Eemeren et al. 2014). By advancing a standpoint, a party assumes a positive or negative position towards a proposition. This commitment to a proposition obliges the party that put forward the standpoint to defend their standpoint (i.e., by advancing supporting argumentation) if challenged to do so by another party (van Eemeren et al. 2014)
Argumentation	Throughout this article, the term argumentation is predominantly used to refer to a collection of arguments. A formal definition of this term, which explicitly contains a process and product dimension, is "a verbal, social and rational activity aimed at convincing a reasonable critic of the acceptability of a standpoint by putting forward a constellation of one or more propositions to justify this standpoint (van Eemeren et al. 2010, p. xii)".
Argument	Understood as simple inferences from a collection of premises to a standpoint. One argument typically consists of two premises: an explicit premise and an implicit premise (or unexpressed premise).
Unexpressed premise	Understood as (intentionally) omitted elements implicitly present in the argumentation for which a party can nevertheless be held accountable (van Eemeren et al. 2010). As such, an unexpressed premise forms the (implicit) link between a (sub)standpoint and one supporting argument. Since argumentation is only logically or pragmatically complete when unexpressed premises are made explicit, parties can be held accountable for an implied unexpressed premise (van Eemeren et al. 2014). (Identified) unexpressed premises may themselves be considered unacceptable or unreasonable.
Difference of opinion	Understood as an (anticipated) disagreement between two parties regarding a given standpoint, i.e., a standpoint can be met with doubt, or an opposing standpoint can be put forward (van Eemeren et al. 2010). Is the driving force for putting forward argumentation, i.e., without any (anticipated) difference of opinion, putting forward argumentation would be pointless (van Eemeren et al. 2014).
Starting point	Set of (typically unexpressed) knowledge, experiences, beliefs, norms and values that provides the basis for standpoints and argumentation put forward in a critical discussion (van Eemeren et al. 2010). Ideally, discussants agree on the starting points of a discussion before commencing in an exchange of argumentation. If there is no shared understanding of the relevant starting points, then misunderstandings may occur or the starting points themselves may become the subject of discussion, rather than the standpoints at issue (van Eemeren et al. 2014). In a scientific discussion, starting points could follow from the particular field of expertise, underlying paradigms and fundamental scientific principles in which a discussant's scholarly work is grounded. Starting points could also follow from an expert's personal beliefs, e.g. related to the roles of experts in policy processes and decision-making and the role of scientific expertise in contemporary society.
Reasonable(ness)	Used to assess the quality of an argumentative move. Argumentation is considered reasonable when it does not contain anything that forms an obstacle to resolving a difference of opinion (van Eemeren et al. 2010). The set of ten discussion rules for a critical discussion can be used to evaluate the reasonableness of argumentation (van Eemeren et al. 2010), though this evaluation is beyond the scope of this article.

overview of the standpoints and all underlying argumentation, and their hierarchal relationships. The three types of argumentation structures are multiple argumentation, coordinative argumentation or subordinative argumentation (see Appendix Table A3.2 for more detailed information). Argument schemes describe the specific type of relationship between a (sub)standpoint and one single underlying premise. The three types of argument schemes are argumentation based on a symptomatic relation, argumentation based on a causal relation or argumentation based on a relation of analogy (see Appendix Table A3.3 for more detailed information). Note that one could also distinguish 'reasoning schemes' (e.g., modus ponens or modus tollens, see Appendix A3.1). However, since reconstruction of such schemes requires that all premises contained in the argumentation are explicit, which is relatively unusual, we rarely use the term 'reasoning schemes'.

3.3 METHOD

Two scientific publications in the field of EDS science have been analyzed using PDAT. These publications are Lamb et al. (2014) and Bergman et al. (2015). Briefly, Lamb et al. (2014) have drafted an elaborate criticism of the WHO-UNEP (2013) report on the state of the science of EDS. Subsequently, Bergman et al. (2015) drafted a rebuttal to criticize the Lamb et al. critique. The 2012 WHO-UNEP report itself is not part of the argumentation analysis. We selected these publications for two reasons. First, both publications discuss a wide variety of aspects of EDS science. Second, the publications present different, sometimes competing argumentation with regard to these aspects of EDS science.

The argumentation analysis of the two publications was performed according to the following steps. First, the main standpoints were identified. Then, argumentation in support of these standpoints was identified and the structure of the argumentation was reconstructed. This yielded two comprehensive argumentation structures, one for each publication. In the case of notable (sub)standpoint—argument relationships, unexpressed premises were made explicit and the apparent argument scheme was identified. To illustrate how these steps work out in practice, a practical example of an argumentation analysis can be found in the Supporting Information. Subsequently, all insights gathered by this argumentation analysis have been used to attempt to make explicit the starting points that we think the two author groups appealed to throughout their publications. Finally, we have identified instances where we think these starting points appear to differ between the two author groups and we identified the type of ambiguity at stake (i.e. primarily interpretative ambiguity, primarily normative ambiguity, or elements of both).

Next, the main standpoint and main supporting argumentation put forward in the two publications will be discussed. Some arbitrarily selected results of our argumentation analysis will be described, since the argumentation structures are too comprehensive to describe in their entirety. These selected examples serve to illustrate how concepts like argumentation structure, argument scheme and unexpressed premise work out in practice, while simultaneously using examples that pertain specifically to EDS science.

3.4 RESULTS

3.4.1 Main elements and specific examples of argumentation put forward in Lamb et al. (2014)

3.4.1.1. Main standpoint of Lamb et al. publication

The main standpoint, which comes forward from our analysis of this publication, is: 'T*he WHO-UNEP 2012 report should not be used to support evidence-based decisions*' (p.37). Although the whole publication focuses on the criticism that the WHO did not provide a balanced perspective and did not provide an update of the state of the science, this does not appear to be the main standpoint. As Lamb et al. consider the 2012 WHO-UNEP report to fall short on some of the aspects that govern such systematically collected and evaluated evidence, one may infer that the report should in fact not be used for evidence-based decision-making. In the concluding paragraph this is explicitly stated (p.37):

Overall, the WHO-UNEP 2012 report on endocrine disruptors fails to achieve its objectives as an updated stateof-the-science review on endocrine disrupting chemicals, and therefore, should not be used to support evidence-based decisions.

The word 'overall' announces a summary of the publication, which starts with one of the main arguments and the standpoint. The indicator 'therefore' shows the argumentation is progressive, which means the subsequent premise could be seen as the main standpoint.

3.4.1.2 Main argumentation structure of Lamb et al. publication

The main argumentation structure of the Lamb et al. publication, as outlined in the argumentation structure below, focuses on the claim that the report was not balanced and did not accurately reflect the state of the science:

1. The WHO-UNEP 2012 report should not be used to support evidence-based decisions
   - 1.1. The WHO-UNEP 2012 report does not provide a balanced perspective
     - 1.1.1 The integrity of decisions at all levels of the 2012 report is questionable
     - 1.1.2 In some instances, the 2012 report failed to consider whether the weight of evidence supports their conclusions or alternative explanations are more likely when they described trends of increasing endocrine-related disorders and concluded these are due to environmental EDCs
     - 1.1.3 The 2012 report does not sufficiently address elements relevant to the definition of EDCs.
     - 1.1.4 The summary for decision-makers has more shortcomings than the report Itself
   - 1.2. The report does not accurately reflect the state of the science on endocrine disruption
     - 1.2.1 The report does not meet the expectations of a state-of-the-science review
     - 1.2.2 The report is not an update of the WHO-IPCS 2002 report

The standpoint is supported by two main arguments (1.1 and 1.2). We consider this argumentation as multiple, because each argument would by itself be sufficient to support the standpoint. For instance, if the report was perceived as balanced but did not reflect the state of the science accurately (according to Lamb et al.), it may still not be considered adequate to support evidence-based decision-making. Similarly, we consider that the four arguments supporting the assertion that the WHO 2012 report did not provide a balanced perspective (1.1) are examples of multiple argumentation. The arguments are independent and provide an alternative defense in case one of the arguments is not accepted by the reader, which is very common in discussions that are directed at a diverse audience. The second main argument (1.2) is also supported by multiple argumentation.

3.4.1.3 Example of complex (subordinative) argumentation

Because of the relative strength of the claim that the integrity of decisions in the report is questionable (1.1.1), this kind of argumentation would need support from further subordinative argumentation:

1.1.1 The integrity of decisions at all levels of the 2012 report is questionable
- 1.1.1.1a The impression has been given that the weight of evidence for causation is stronger than it is
  - 1.1.1.1a.1a Unjustified inferences are made to suggest causation
  - 1.1.1.1a.1b Literature was cited selectively, without discussion of contradictory studies
- 1.1.1.1b Conclusions were predisposed to the identification of potential EDCs
  - 1.1.1.1b.1 No adequate systematic approach has been used to assess causation

According to our reconstructed argumentation structure, substandpoint 1.1.1 is supported by argument 1.1.1.1a, which refers to an impression of misrepresentation of the weight of evidence. As impressions typically have limited persuasive force and thus generally need further explanation, this argumentation is supported by more additional subordinative argumentation.

3.4.1.4 Example of identifying and evaluating an unexpressed premise

In some cases, the unexpressed premise may give new information that has consequences for the argumentation, as PDAT proposes that the protagonist should also be held committed to premises left implicit. If these premises are not made explicit, these commitments may be overlooked or a potential weaker link in the argumentation may remain ignored. The unexpressed premises in the article written by Lamb et al. mostly contain premises that have elsewhere also been stated explicitly. The following example deals with argumentation concerning adequate approaches to causation. Usually, making unexpressed premises explicit takes two steps. First, an attempt is made to reconstruct the logical reasoning. This involves adding a premise that makes the argument logically valid. Second, the unexpressed premise is made more informative. We consider that the substandpoint that no adequate systematic approach has been used to assess causation (1.1.1.1b.1) is supported by the argument that

Bradford Hill's criteria (or similar systematic methods to assess causation) have not been used, (1.1.1.1b.1.1a). The (formal) reasoning can
be reconstructed as follows, using modus ponens as the basic reasoning scheme:

1. If p, then q (If Bradford Hill's criteria (or similar systematic methods) have not been used, then no adequate systematic approach has been used to assess causation.)
2. p (Bradford Hill's criteria (or similar systematic methods) have not been used.)
3. q (No adequate systematic approach has been used to assess causation.)

To make this modus ponens structure more informative, the missing premise could be made explicit:

(1.1.1.1b.1.1a') Bradford Hill's criteria (or similar systematic methods) are an adequate systematic approach to assess causation

In this case, the unexpressed premise shows that there is an assumption, or starting point, that Bradford Hill's criteria are in fact considered adequate to assess causation systematically. Thus, making unexpressed premises explicit may provide information about starting points and subsequently bring to light a difference in starting points. Other stakeholders, such as the authors of the WHO-UNEP 2012 report, may not hold the same opinion regarding this starting point.

3.4.2 Main elements and specific examples of argumentation put forward in Bergman et al. (2015)

3.4.2.1 Main standpoint of Bergman et al. publication

The main standpoint of the response to the critique of "State of the Science of Endocrine Disrupting Chemicals 2012" coming forward from our analysis of this publication is: 'The criticism by Lamb et al. on the 2013 WHO-UNEP report is unjustified'. The standpoint could be detected quite easily. First, we interpret the response as a defense against the critique that the report should not be used to support evidence-based decisions, so the authors of the rebuttal will presumably refute this criticism. Second, already in the abstract, the Bergman et al. publication appears to presents various reasons why, in their opinion, the critique was flawed. That is, according to our interpretation of the Bergman et al. publication, Lamb et al. have quoted the 2012 WHO/UNEP report in an incomplete and misleading fashion, misused conceptual frameworks for assessing causality and defined extremely narrow standards for synthesizing and reviewing evidence, among others. Third, according to Bergman et al., the authors of the critique have not directed their messages at the scientific community but at decision makers instead. This implies that the critique lacks a focus on scientific issues, which has also been stated explicitly (p. 1016).

3.4.2.2 Main argumentation structure of Bergman et al. Publication

Our reconstruction of the main argumentation structure of the Bergman et al. publication is shown below:

1. The criticism by Lamb et al. on the 2013 report is unjustified.
  1.1 It created the false impression of a scientific controversy
    1.1.1 Lamb et al. falsely claimed that the integrity of decisions at all levels of the WHO report should be called into question
      1.1.1.1 Lamb et al. had alternative motives for writing this critique: to confuse the scientific data
        1.1.1.1.1 They misdirected the reader into thinking that the 2012 report was biased
        1.1.1.1.2 They employed the same tactics as the tobacco industry to undermine attempts of introducing standardized packaging for cigarettes
        1.1.1.1.3 They were sponsored by the chemical industry
  1.2 It does not engage with the scientific substance of the report
    1.2.1 The claim that the report is neither a state of the science nor an update of the previous report is false
    1.2.2 The claim that the lack of a formal assessment of causation and weight-of-evidence approaches leads to subjectivity and the suggestion of causation is false
    1.2.3 Lamb et al.'s claim that other environmental causes of disease trends than chemicals were not acknowledged is false
    1.2.4 Lamb et al.'s criticism on our [Bergman et al.'s] characterization of the endocrine system is unfounded
    1.2.5 The criticism against the summary for decision-makers is unfounded

The main arguments for the standpoint that the criticism is unjustified are that it creates the false impression of a scientific controversy and that the critique does not engage with the scientific substance of the report. This already indicates that the critique was perceived as misleading by Bergman et al., in the sense that the critique supposedly ignores the science of endocrine disruptors and tries to create a scientific controversy that did not exist. The latter is motivated by attributing alternative motives onto Lamb et al. as an explanation of why this alleged 'artificial controversy' was constructed. We consider all arguments as independent, since they provide alternative reasons for concluding that the authors had alternative motives. Each of them would be sufficient to prove that the standpoint is true (assuming that the premises are true). The claim that the critique did not engage with the scientific substance of the report is supported by various examples of how the critique misconstrued the report and did not accurately reflect the scientific discussions under scrutiny.

3.4.2.3 Example of identifying and evaluating an argument scheme

The following example shows how an unexpressed premise helps to determine what argument scheme has been applied.

1.1.1.1 Lamb et al. had alternative motives for writing this critique: to confuse the scientific data
1.1.1.1.1 They misdirected the reader into thinking it was biased
1.1.1.1.2 They employed the same tactics as the tobacco industry to undermine attempts of introducing standardized packaging for cigarettes
1.1.1.1.3 They were sponsored by the chemical industry

The misdirection and tactics that were used to present the WHO-UNEP 2012 report as unreliable, as well as the connections to the chemicals industry, are seen as evidence for the statement that Lamb et al. had alternative motives for writing this critique. Without the unexpressed premise, though, it is not totally clear why the second argument would be relevant. If the analogy was a comparison based on similar tactics, it is unclear what the relevance would be of mentioning the tobacco industry and in what way this would help prove the statement to be acceptable. The fact that the tobacco industry was mentioned is the point that actually creates the weight of the argument. By making the unexpressed premise explicit, it becomes clear that the argument scheme used here is argumentation based on a relation of analogy, since an analogy has been made between the tactics used by the chemical industry and those used by the tobacco industry:

(1.1.1.1.2') And the tactics used by the tobacco industry are comparable to the tactics used by the chemical industry

As Lamb et al. acknowledge that in drafting their publication, they received funding support from several chemical industry sponsors, their argumentation is seen as representative of this branch of industry by Bergman et al. (as has been made explicit with argument 1.1.1.1.3). The proposed similarity of tactics employed by the tobacco industry and alleged tactics employed by the chemical industry is what appears to have led to the attribution of alternative motives onto Lamb et al. The argument scheme based on a relation of analogy is as follows:

Y is true of X (having alternative motives is true of Lamb et al.)
because: Y is true of Z (having alternative motives is true of the tobacco industry)
and: Z is comparable to X (the tactics used by the tobacco industry are comparable to the tactics used by Lamb et al./the chemical industry)
To ensure that the analogy is a sound one, the most important critical question to ask is:

Are there any significant differences between Z and X?

Naturally, one is always able to find differences between Z and X, because if they were exactly the same, making an analogy would be pointless and an example of circular reasoning. The point is that they should be the same concerning all characteristics that are relevant to the argument (as discussed above).

3.4.3 Differences in starting points

In this section, we concentrate on the part of the analysis that we consider crucial in understanding why it will be very difficult to reach a satisfactory concluding stage in the discussion between the two author groups. The differences in starting points that appear to impede the path to resolution are addressed in more depth. Table 3.2 gives an overview of differences in starting points we identified, based on a close evaluation of the argumentation structures of Lamb et al. (2014) and Bergman et al. (2015), including notable argument schemes and unexpressed premises.

The first difference in starting points we identified revolves around the question whether this discussion falls into the category of a scientific controversy. Bergman et al. have stated explicitly that Lamb et al. create the false impression of a scientific controversy to confuse the scientific data. By saying this, Bergman et al. do not appear to agree with the statement of Lamb et al. that it is a scientific controversy. Alternatively, Lamb et al. have implied that the report by the WHO creates a false sense of agreement in some instances, which indicates that they hold the opinion that on important matters there is no agreement, which in turn appears to imply that they do believe there is a scientific controversy.

The second difference in starting points we identified revolves around the preference for a specific weight-of-evidence approach. Lamb et al. have implied that an objective and structured weight-of-evidence approach for EDS exists, whilst Bergman et al. have stated explicitly that such an approach does not (yet) exist for EDS. Lamb et al. mention a series of requirements that make up for such an objective weight-of-evidence approach throughout their publication. However, Bergman et al. argue that objective weight-of-evidence approaches do not exist. For example, hypothesis formation is considered to be an inherently interpretative process that cannot be dealt with in an 'objective' manner. In this light, the approach deemed most appropriate is the "best professional judgment".

The third difference in starting points we identified concerns the existence of a systematic approach to assess whether an association is causal and what this approach should entail. Apart from the belief that such an approach exists, Lamb et al. describe various conditions that have to be fulfilled before the criteria for causality will be met, predominantly inspired by Bradford Hill's criteria. Nevertheless, Bergman et al. give the impression that Lamb et al. have unrealistic demands, as such an approach is not deemed feasible when proof of causality in the field of EDS science is supposedly very difficult, if not impossible. They assert that no approach will completely protect against bias and that absolute criteria to assess causality are impossible. Therefore, a different approach is deemed unavoidable.

The fourth difference in starting points we identified pertains to a disagreement about assumptions held concerning the physiological function of the endocrine system and its potential to cope with transient exposures to environmental chemicals. Lamb et al. appear to portray the endocrine system as a homeostatic system adaptable to circumstances. Alternatively, Bergman et al. appear to portray the endocrine system as vulnerable to

Table 3.2: Overview of the identified differences in starting points.

Nr.	Topic	Lamb et al.	Bergman et al.	Type of ambiguity
1	Degree of controversy within scientific EDS debate	The 2012 WHO-UNEP report did not acknowledge controversy over the interpretation of data in several instances (creating a false sense of agreement)	This is not a scientific controversy; Lamb et al create the fake impression of scientific controversy	Primarily interpretative ambiguity
2	Weight-of-evidence approaches	Objective methods to evaluate the weight of the evidence do exist; a 'best professional judgment' approach is not warranted for a state-of-the-science report and prone to bias	Objective methods to evaluate the weight of the evidence do not exist; scientific judgments are inherently required, warranting the 'best professional judgment' approach	Elements of both interpretative ambiguity and normative ambiguity
3	Establishing causality and use of Bradford Hill's criteria/ viewpoints	Bradford Hill's criteria are an adequate starting point to unequivocally establish causation	Bradford Hill's viewpoints cannot be applied unequivocally; e.g. the social, economic or political context may influence the required strength of evidence required to take policy action	Primarily normative ambiguity
4	Framing of endocrine system	The endocrine system is resilient, as it is specifically designed to cope with environmental chemical exposures, through natural homeostatic processes (that is, within the boundaries of homeostasis)	The endocrine system is vulnerable to Environmental exposures. The endocrine system of developing fetuses is susceptible to transient fluctuations of circulating hormone, potentially leading to irreversible malformations	Elements of both interpretative and normative ambiguity
5	Function of a state-of-the-science report	Robust, objective and systematic scientific procedures are central to a state-of-the-science report that is to be used to support evidence-based decision-making	The underlying scientific standards of a state-of-the-science report should be explicitly responsive to considerations of public health protection	Primarily normative ambiguity

irreversible disruption, mainly due to its programming functions in developing (unborn) children. These differing illustrations of the endocrine system appear to lead to equally differing judgments about how disturbing an alteration of the endocrine system may be at a given life stage of the exposed individual.

The fifth difference in starting points is more fundamental and shows overlap with some of the differences in starting points described above. In essence, it appears that Lamb et al. and Bergman et al. have differing perspectives regarding the function of a state-of-the-science report, particularly in the context of how evidence-based decision-making should be supported. On the one hand, Lamb et al. appear to emphasize the necessity of a state-of-the-science report to live up to their proposed scientific standards and approaches. The authors note that, only in this way, one can then be confident in the decision-making that arises from such a state-of-the-science report. On the other hand, Bergman et al. appear to emphasize that the underlying scientific standards of a state-of-the-science report should be explicitly responsive to considerations of public health protection. From the perspective of Bergman et al., the level of evidence or strength of association that is necessary to take policy action cannot be determined by some 'fixed' standard, but rather depends on the decision-making context. For example, under certain circumstances, a weak strength of association may be very significant from a public health perspective.

3.5 DISCUSSION

In this study, we have analyzed two scientific publications in the debate on EDS science from an argumentation point of view, using the pragma-dialectical argumentation theory. From this argumentation analysis, five differences in starting points have been identified. These differences in starting points will be characterized based on the nature of the disagreement. Where relevant, similarities to existing risk and expert role typologies from scientific literature will be identified. Finally, some strengths and limitations of this argumentation analysis will be discussed.

3.5.1 Starting points and types of ambiguity

We draw from Renn (2008) to discern interpretative ambiguity and normative ambiguity and subsequently distinguish between the type of value differences at stake. We consider the first difference in starting points (degree of scientific controversy) in Table 3.2 as primarily interpretative ambiguity, while the third (establishing causality) and fifth (function of a state-of-the-science report) difference in starting points are considered to be primarily normative ambiguity. The second (weight-of-evidence approaches) and fourth (framing of the endocrine system) difference in starting point share elements of both interpretative and normative ambiguity.

3.5.2 Analogies of starting points with existing risk and expert roles typologies

The fourth difference in starting points pertains to differing perspectives of the function of the endocrine system in the physiology of humans and wildlife. This fourth difference in starting point may, at the human physiological level, not be conflicting, such that both perspectives can

co-exist. We are interested, however, in why the two author groups appeared to elect one position over the other in the argumentation presented in their publications. We argue that this specific difference in starting points could be explained by differences in culturally determined perspectives of risk. To illustrate this, we use the concept of 'Myths of Nature' (Dake 1992; Steg and Sievers 2000), derived from the 'Cultural Theory of Risk' (Douglas 1992; Rayner 1992; Tansey and O'Riordan 1999). Dake (1992, p. 24) defines a Myth of Nature as 'one set of beliefs about what the world is like, what its risks are like, and who is to blame for untoward events'. A Myth of Nature is considered a specific type of cultural bias, which is a set of values and beliefs shared within a group (Dake 1992). Indeed, a Myth of Nature is a cultural bias that specifically pertains to environmental risks. According to Cultural Theorists, individuals draw from their cultural biases to conceive their personal perceptions and interpretations of a risk (see e.g. Rayner 1992). Insight in one's Myth of Nature may then explain why an individual may emphasize some aspects of a risk, such as the risk's potential hazardous impacts to human health, whereas other aspects may remain underexposed, such as the benefits associated with the risk and the potential lack of viable or marketable alternatives. Cultural Theorists generally discern four prototypical Myths of Nature: 'nature is capricious', 'nature is tolerant, but within limits', 'nature is benign' and 'nature is ephemeral or fragile' (Dake 1992) (see Table 3.3; Figure 3.1).

On the basis of the starting points identified from our argumentation analysis, we propose that Lamb et al.'s illustration of the endocrine system in human physiology resembles the 'nature is tolerant, but within limits' Myth of Nature. Lamb et al. note that a key function of the endocrine system is to deal with continuous fluctuations in hormone levels, though the limits of these homeostatic processes need to be respected. This perspective is much similar to the perspective of nature as a tolerant system that has its boundaries. Alternatively, we propose that Bergman et al.'s illustration of the endocrine system resembles the 'nature ephemeral or fragile' Myth of

Table 3.3: Four Myths of Nature and their key characteristics (based on Dake, 1992 and Steg and Sievers, 2000).

Myth of nature	**Key characteristics**
Nature is capricious	Nature is considered to be an unmanageable system. Supporters of this cultural bias are hypothesized to believe that incidents and accidents happen by chance and cannot be predicted
Nature is tolerant, but within limits	Nature is considered a robust system, but which has its boundaries. Supporters of this cultural bias are hypothesized to believe that the appropriate authority (i.e. experts and decision makers) can derive and establish the limits of nature
Nature is benign	Nature is considered a robust and resilient system. Supporters of this cultural bias are hypothesized to believe that natural systems have the inherent capability to cope with virtually any (man-made) impacts
Nature is ephemeral or fragile	Nature is considered a fragile system. Supporters of this cultural bias are hypothesized to believe that only minor disturbances of a natural system could lead to catastrophic, irreversible consequences

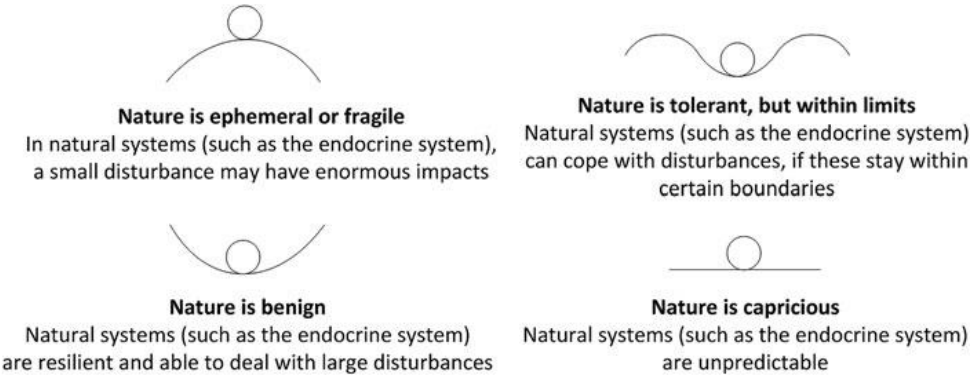

Figure 3.1: The four Myths of Nature, represented by a ball (behavior associated with risk) and a landscape (the type of natural system).

Nature. Bergman et al. emphasize the susceptibility of vulnerable groups to irreversible disruption of the endocrine system leading to potentially latent effects, even in situations of (very) low exposures to EDS. As such, the Myth of Nature typology draws attention to differences in beliefs about the endocrine system and subsequent differences in framing of this system. Whether this typology has added value in understanding the positions of experts in the broader discussion on EDS science should follow from additional research.

The fifth difference in starting points pertains to different emphases with regard to the function of a state-of-the-science report. A recurring theme in the Lamb et al. publication appears to be the scientific standards to which a state-of-the-science report on EDS science should live up to. For example, a summary is provided at the end of the Lamb et al. publication that pinpoints why the authors think the 2012 WHO-UNEP report cannot be characterized as a state-of-the-science review:

> *"[due to] the lack of a defined scope for the review, the absence of a process for identification, integration and interpretation of data, the lack of a structure for evaluating individual studies for relevance and reliability, and an objective method for evaluating the weight of the evidence" (Lamb et al. 2014, p. 36).*

In this summary, the various reasons put forward by Lamb et al. are all grounded in characteristics of what they appear to consider robust science. The merits of a state-of-the-science report therefore appear to be determined by its ability to live up to the scientific standards comprehensively described in the Lamb et al. publication. Notably, this does not mean that Lamb et al. may have little concern for human health and the environment, but rather that objective and structured approaches to analyze the available evidence are repeatedly emphasized. In the Bergman et al. publication, extensive substantiation of the scientific procedures used to develop the 2012 report are provided (though different standards and

approaches are applied than those advocated Lamb et al., see e.g. the second and third differences in starting points). However, in addition to a discussion of what they consider appropriate scientific approaches, Bergman et al. refer multiple times to the importance for a state-of-the-science review (and the toxicological and epidemiological sciences in a broader sense) to have utility for public health initiatives. This point is addressed explicitly in their publication, while simultaneously alleging Lamb et al. to overlook this perspective:

> *"Lamb et al. are remiss in acknowledging that the goal of the toxicological and epidemiological sciences is not to provide assessments as an end in themselves, but to explore and evaluate conditions that offer disease prevention and public health initiatives" (Bergman et al. 2015, p. 1011).*

Among others, Bergman et al. describe that even a weak strength of association could very well provide grounds for protective measures, or at least does not exempt one to consider such measures, since public health may be served significantly. In addition, they describe that hypotheses formulated in the state-of-the-science review must have some utility in serving public health considerations.

Notably, this particular difference in starting points shows similarities to a theoretical typology developed by Weiss (2003). This typology's point of departure is the observation that the various formulations of the Precautionary Principle generally do not accurately define the required level of scientific certainty to justify a particular precautionary measure, such as 'measures against most serious aspects' (Weiss 2003). Note that, while some may question the use of the term 'level of (scientific) uncertainty', we prefer to use this term in this context as it is consistently used by Weiss and this term is a key aspect of the reasoning underlying the Weiss (2003) typology.

The typology of Weiss (2003) can be used to characterize the roles of experts in terms of their stance towards the Precautionary Principle. Five ideal-typical expert roles can be discerned (see Figure 2.3 in Chapter 2). The scientific absolutist, typically being an advocate of 'science-based regulation', will require a high level of scientific certainty before supporting 'measures against the most serious aspects' of a new technology that may pose potential dangers to the environment. The environmental absolutist, typically being an advocate of relatively early precautionary action, will require much less scientific certainty to advocate the same type of measure, supported by the norm that the environment inherently requires protection from potential (man-made) dangers. The cautious environmentalist, environmental centrist and technological optimist hold intermediate positions.

Whether the expert role typologies of Weiss (2003) do in fact accurately capture the identified difference in starting points needs further research, since both of the analyzed publications make little explicit references to preferred policy measures regarding particular EDS. The typology does draw attention to differences in the roles that experts could adopt to provide

science-based advice to policy makers. Spruijt et al. have empirically identified, among others based on the theoretical insights from Weiss (2003), differences in expert advisory roles in the field of electromagnetic fields risk research (Spruijt et al. 2015) and particulate matter risk research (Spruijt et al. 2016). Whether these differences in expert roles would also be visible in the field of EDS science would similarly be a topic for future research.

3.5.3 Strengths and limitations

This is the first argumentation analysis using PDAT to study (a part of) the scientific discussion on EDS, as far as we are aware. This study combined knowledge from the field of EDS science with insights from argumentation theory, values in science and existing theoretical risk and expert role typologies, making this an interdisciplinary effort. Our main analytical framework, the pragma-dialectical argumentation theory (PDAT), allowed us to systematically analyze the argumentation put forward in Lamb et al. (2014) and Bergman et al. (2015) in its entirety. In turn, this approach allowed us to focus on assumptions that can play a major role in an exchange of argumentation, but nevertheless remain implicit and, thus, somewhat intangible. The elaborate methodology of PDAT should ensure that such argumentation analyses remain true to the essence of the text(s) under scrutiny. We argue that combining the systematic analytical framework offered by PDAT with (a) insights from literature studying values in science (see e.g. Douglas 2000; Elliott 2017) and (b) existing typologies of risk and expert roles have provided us with specific 'lenses' to study pertinent value differences. This allowed us to distinguish between interpretative and normative ambiguity, and to ultimately identify various types of value differences occurring in a specific part of regulatory science on EDS. Moreover, we identified several differences over normative values that, at first glance, appeared to be camouflaged as conflicts of scientific fact.

Only two publications from one period in time have been analyzed, while the literature on endocrine disruptor science is far more extensive. The analyzed publications are embedded within an expansive, on-going scientific discussion that may cover a wider range of topics than discussed in the particular publications analyzed in this study. As this research remains an exploration of the viability of argumentation analysis to identify some apparent value differences in the scientific discussion on EDS risks, we argue that analyzing two publications was justified. Similarly, since PDAT, our analytical method of choice, requires the analyst to necessarily limit the analysis to the specific substance of the arguments at hand, the scope of our analysis is limited accordingly. However, we think that our specific research aims warrant the limited scope of our analysis.

Finally, we did not evaluate potential fallacious reasoning or provide any answer to critical questions raised by our argumentation analysis. We aimed to analyze the scientific EDS discussion from some distance to perform an analysis that is as unbiased and impartial as reasonably possible. Evaluations of fallacious reasoning would necessarily require us to make normative judgments, which would compromise our neutral position and may potentially influence our credibility as (argumentation) analysts.

3.6 CONCLUSIONS

In this article, we set out to show how an argumentation analysis can be used to identify the main standpoints, main supporting arguments and other argumentation. We have also shown that such an approach can be used to highlight differences in starting points and subdivide contrasting starting points based on the type of ambiguity, and subsequently the type of value differences at issue. We analyzed two scientific publications in the field of EDS science to demonstrate this approach. Our results show a collection of five differences in starting points, two of which have been further investigated using existing risk and expert role typologies. We argue that it would be ineffective for at least four of these five differences in starting points to attempt to settle these solely in the scientific sphere, due to the nature of the (often ethical) value differences involved. Rather, multi-stakeholder approaches are then required. Such approaches would in practice be well served by further argumentation analysis of the pertinent value-laden positions involved. Future research could show to what extent the differences in starting points identified in this article can be extrapolated to the broader discussion on EDS science, or even to scientific discussions on other environmental health risk controversies. Finally, additional research could use the risk and expert role typologies applied in this article to further explore the roles of EDS experts when supporting evidence-based decision-making.

ACKNOWLEDGMENTS

The authors would like to thank Ric van Poll, Tom Jansen and two anonymous reviewers for their insightful comments on earlier versions of this article.

Analysis of different preferences for the EU's regulatory options for endocrine disruptor identification criteria using argumentation theory

Sander C.S. Clahsen, Lana Moss, Irene van Kamp, Theo G. Vermeire, Bart J. Garssen, Aldert H. Piersma, Erik Lebret

Published in:
Science of the Total Environment
Volume 740, 2020

ABSTRACT *What criteria are most suitable to identify endocrine disrupting substances (EDSs) for regulatory purposes in the EU? The results of the European Commission's public consultation, as part of the process to establish identification criteria for EDSs, show that different regulatory options are supported. Some respondents prefer an option including hazard characterization considerations, whereas others prefer an option that avoids these considerations and introduces several hazard-identification based weight-of-evidence categories. In this study, the argumentation underlying the different preferences for identification criteria are analyzed and compared using pragma-dialectical argumentation theory (PDAT). All responses of non-anonymous, national governments that submitted a response in English (n=17) were included. Responses of other stakeholder organizations were included if a Google News search returned an opinionated presence in the media on the subject (n=9). Five topical themes and 21 underlying issues were identified. The themes are 1) mechanistic understanding of EDSs, 2) regulatory considerations related to the identification of EDSs, 3) consistency with existing regulatory frameworks, and 4) evaluations of specific issues related to a category approach and 5) related to including potency. We argue that two overarching (implicit) 'advocacy coalitions' can be discerned, that adopted contrasting positions towards the identified themes and issues. Among these 'coalitions', there appears to be consensus about the necessity of having 'science-based' criteria, though different perspectives exist as to what the most accurate mechanistic understanding of EDSs entails. To move the discussion forward, we argue that a societal dialogue would be beneficial, where EDS science and regulation are discussed as interrelated themes.*

4.1 INTRODUCTION

For several years, the European Commission (EC) has been working towards the regulation of endocrine-disrupting substances (EDSs), most notably in the areas of pesticides and biocides. The EU's Plant Protection Products Regulation (PPPR, No. 1107/2009) required the Commission to adopt identification criteria to establish whether a pesticide active substance should be considered an endocrine disrupting substance (EDS). The EU's Biocidal Products Regulation (BPR, No. 528/2012) contains a similar requirement for such criteria. In the 2014 Roadmap of the Commission (EC, 2014), four options for criteria were presented. These options mainly differ as to the weight of evidence necessary to identify an EDS as such (see Table 4.1 for a description of these options, and the key criteria).

As part of the selection of the criteria, an impact assessment was performed, which also included a public consultation. In this consultation, information was requested about the various potential impacts of the four options for criteria: the range of substances that could be identified under each option, the potential for substitutability of these identified substances and anticipated socio-economic impacts, among other aspects (EC, 2015). Notably, there was also room to provide general comments, which was used as an opportunity by many respondents to state which option they preferred, and which they opposed, along with supporting argumentation.

The report on the results of the public consultation (EC, 2015) has shown that there are different perspectives among respondents about what the ultimate criteria should entail. The main aim of the present paper is to analyze and compare the argumentation underlying different option preferences of governmental entities (e.g. national governments) and of prominent stakeholder organizations (e.g. NGO's or industry organizations), as stated in their responses.

We were particularly interested in the debate about the EDS identification criteria, because it is characterized by ongoing controversy, both in terms of science and policy. There is general agreement on the scientific definition of an EDS, proposed by the WHO (2002): '*an endocrine disruptor is an exogenous substance or mixture that alters function(s) of the endocrine system and consequently causes adverse health effects in an intact organism, or its progeny, or (sub)populations*'. Accordingly, this definition includes three key elements: 1) exposure to the exogenous substance should cause an adverse health effect; 2) disruption of the function of the endocrine system should be the mechanism of action; and 3) there should be a causal relationship between the exposure, the mechanism of action and the adverse effect. However, practical use of this definition and the key elements has led to much debate. Clahsen, van Klaveren et al. (2020) have shown that there are fundamental differences of opinion between EDS experts in how weight-of-evidence evaluations for EDS should be performed, and whether there are systematic approaches available and useful for establishing causality, among other aspects. These elements are crucial for the development of sound science-informed methodologies for the identification of EDSs, irrespective of the ultimately selected option.

Table 4.1: Description of the four policy options proposed by the EC, including the key identification criteria. Bold phrases highlight the key elements of that option.

Policy option	**Description based on information provided in EC (2014)**	**Key criteria to identify a substance as EDS**
Option 1 ("Interim criteria")	The **interim criteria** as included in the PPPR and BPR will remain in place (see point 3.6.5. in the legal text of the PPPR and Article 5.3 in the legal text of the BPR).	> Carcinogenic category 2 and Reproductive toxicant category 2 under the EU's classification, labelling and packaging (CLP) regulation (EC, 2008) for carcinogenic, mutagenic and reproductive toxic (CMR) substances > Reproductive toxicant cat. 2 under CLP and occurrence of toxic effect to endocrine organ (note: a substance *may* then be identified as EDS)
Option 2 ("WHO/IPCS definition")	The **WHO/IPCS definition** of an endocrine disruptor will be used to identify EDSs, in combination with several weight-of-evidence requirements (see EC, 2014 for more information).	> Exposure to the substance causes an adverse health effect > As mechanism of action, disruption of the function the endocrine system is identified > There is a causal relationship between the mechanisms of action and the adverse effect
Option 3 ("Category approach")	The WHO/IPCS definition of an EDS, including the outlined weight-of-evidence requirements, will be used as a basis to identify EDS, but **additional categories** will be included that refer to different strengths of evidence. Next to category I (**'endocrine disruptor'**), which is equivalent to option 2, the categories II (**'suspected endocrine disruptors'**) and III (**'endocrine active substances'**) are added. The specific weight-of-evidence requirements related to the two additional categories can be found in EC (2014).	> Category I: see criteria for Option 2 above. > Category II: substances could be allocated to this category on the basis of some evidence for endocrine-mediated adverse effects, but which is not sufficient to warrant placement in Category I > Category III: substances could be allocated to this category on the basis of some in vitro or in vivo evidence indicating a potential for endocrine mediated adverse effects
Option 4 ("Potency inclusion")	The WHO/IPCS definition of an endocrine disruptor will be used, but a **potency threshold** will be included to discriminate between high potency and low potency EDS. Further information on weight-of-evidence requirements is not specifically mentioned.	> The same criteria as applying to Option 2, but with the inclusion of potency as an element of hazard characterization

The complexity of EDS policy is related to the high stakes surrounding the practical uses of EDSs. Many substances linked to endocrine disruptive effects, such as the plastics constituent Bisphenol A or certain pesticides, are high production volume substances that have wide-spread applications in contemporary society. Some stakeholders argue that the availability and use of some (potential) EDSs could be associated with substantial economic value, both directly and indirectly (see e.g. ADAS, 2011; PlasticsEurope, 2019). Others refer to the significant potential health impacts and related economic costs (see e.g. Rijk, van Duursen and van den Berg, 2016; Norden, 2014). Amidst these stakes, the lobby of the chemical industry has been accused of deliberately obstructing regulatory action against EDSs (Horel and CEO, 2015), while there are also EDS experts that question the motives of NGOs for deliberately maintaining the issue of EDS on the public and legislative agenda (Dietrich et al., 2016).

To ensure a maximally systematic, unbiased and impartial analysis of the argumentation put forward in the selected responses, we used pragma-dialectical argumentation theory (PDAT). With PDAT, the analyst can identify the standpoints and the underlying structure of argumentation put forward. When the same or strongly similar arguments, or clusters of arguments, are repeatedly occurring in multiple responses, this points to the presence of important topics in the debate on EDS identification criteria. In this study, we distinguish two levels of such topics: broad, topical *themes* (e.g. about the broad mechanistic understanding of EDSs) and underlying *issues* touching upon specific aspects of a *theme* (e.g. about the role of timing of exposure). When multiple perspectives towards a *theme* have been observed, these are referred to as contrasting *positions*.

Based on our earlier work in Clahsen, van Kamp et al. (2019), we discern different dynamics, social stations and underlying drivers of argumentation in science-policy controversies, as illustrated in Figure 2.12. From this framework, we derived two pertinent topics of interest for our analysis. First, we are particularly interested in distinguishing science-based argumentation from normative value judgments. This notion provides the starting point for the identification of *themes* and thereby focused the scope of our analysis. That is, the main focus is on the intrinsic properties of the regulatory options, and how these relate to existing perspectives on EDS science and regulation, rather than on discussing arguments pertaining to the potential consequences of implementing one of the four policy options. Second, we study the alignment of arguments in implicit or explicit 'advocacy coalitions' (after Sabatier and Jenkins-Smith, 1993). The use of the concept of 'advocacy coalitions' serves as a heuristic to delineate groups of actors that share the same policy preferences, use similar supportive arguments by referring to the same *themes* and *issues*, and adopt the same *positions* towards these *themes* and *issues*. Note that advocacy coalitions, in the true meaning of the concept, cannot be identified in this study, since actual interactions within and between coalitions cannot be studied. Specific attention will be given to the distribution of governmental entities and stakeholder organizations over these 'coalitions'.

We specified five research questions from the aim of our study:

1. What types of option preferences exist among the identified responses?
2. What arguments have been put forward in favor of or against the four regulatory options for identification criteria (see Table 4.1)?
3. What topical *themes*, underlying *issues*, and contrasting *positions* towards these *themes* and *issues* can be derived from the range of arguments identified?

4a. To what extent can (implicit) advocacy coalitions be identified?

4b. How are governmental entities and prominent stakeholder divided over these advocacy coalitions?

Research question 1 is addressed by a document analysis. Research question 2 is addressed by the analysis of arguments using PDAT, while research questions 3, 4a and 4b are addressed by the subsequent categorization of arguments.

4.2 METHODS

4.2.1 Pragma-dialectical argumentation theory (PDAT)

PDAT was developed by van Eemeren and colleagues (van Eemeren and Grootendorst, 1984; see also van Eemeren, Grootendorst and Henkemans, 2010) to remedy an experienced lack of a systematic way to study argumentation in different social contexts, such as argumentation in scientific discussions or argumentation in 'daily life'. In pragma-dialectics, argumentation is viewed as aiming at resolving a *difference of opinion* by critically testing the acceptability of the standpoints at issue.

PDAT provides a model of argumentation that enables an analysis and evaluation of argumentation. A full analysis of the argumentation provides an analytic overview consisting of a characterization of the difference of opinion, the standpoints, the discussion stages, the *argumentation structure* and the *argument schemes*. An *argumentation structure* provides a complete overview of standpoints and all the underlying argumentation, including hierarchal relationships. PDAT distinguishes three types of *argumentation structures*: *multiple argumentation*, *coordinative argumentation* and *subordinative argumentation*. *Argument schemes* provide the specific relationships between individual arguments. Since the aim of this study is to identify and categorize the range of arguments, rather than analyzing the (logical) links between individual arguments, *argument schemes* have not been identified here. Furthermore, from the perspective of PDAT, it is not appropriate to use the concept of *difference of opinion* in the context of this study. There is no explicit argumentative exchange between the various respondents, since all responses are aimed at the Commission. In our analyses, we use the terms *themes, issues* and *positions* to refer to topical and broad (*themes*) or more specific topics (*issues*) that appear to be under discussion, and the often contrasting *positions* towards the identified *themes* and *issues*, respectively.

4.2.2 Selection of responses included in the argumentation analyses

For our analysis, we used the publicly available database of the Commission. The public consultation elicited 27.087 responses in total. Figure 4.1 shows our selection procedure. The Commission only made public non-confidential and non-email responses (n=22.269). Responses of individuals were excluded, leaving 818 responses of affiliated responses. From these, we selected responses of governmental entities and of prominent stakeholder organizations.

4.2.2.1 Selection of governmental entities

We identified responses from 19 national governmental entities, based on their selected identification as a 'Public authority', excluding one levy board, three anonymous responses, four non-English responses and three local governments' responses.

4.2.2.2 Selection of stakeholder organizations

The public consultation questionnaire generated responses from 788 non-governmental stakeholder organizations, a number too large to subject to argumentation analysis. We therefore focused on those stakeholder organizations that were the most prominent in the societal and political debate on the EDS identification criteria. Accordingly, we performed an online search of news media outlets using Google News (date of search: 26 February 2018) to identify stakeholder organizations that have an 'opinionated' presence in the debates. Stakeholder selection criteria were 1) inclusion only when an explicit opinion was provided to any of the four proposed options in at least one of the selected news articles 2) exclusion of governmental entities and 3) exclusion of professionals or experts providing their opinion on a personal basis. We identified 18 stakeholder organizations on the basis of these selection criteria, nine of which participated in the public consultation and were included in our analysis.

4.2.3 Identification of the preferred policy option in each response

The type of preferred policy option was identified by selecting the key sentence or phrase that appeared to most explicitly convey the preferred policy option in each response. Paragraph A4.1 in the appendix substantiates this procedure, by presenting the key sentence or phrase that we used to discern the range of preferred policy options.

4.2.4 Argumentation analyses

PDAT was used to reconstruct the argumentation structure for the selected responses. These argumentation structures consist of all standpoints and any underlying *single*, *multiple*, *coordinative* or *subordinative* argumentation that the analysts have encountered in the responses, including the hierarchal relationships.

To reduce the influence of possible personal bias and style of analysis on reconstructing argumentation structures, the argumentation structure of one response (ECPA) was reconstructed independently by two authors (LM and SCSC). The resulting minor differences were discussed with a third author (BG), who provided recommendations for the analysis of the

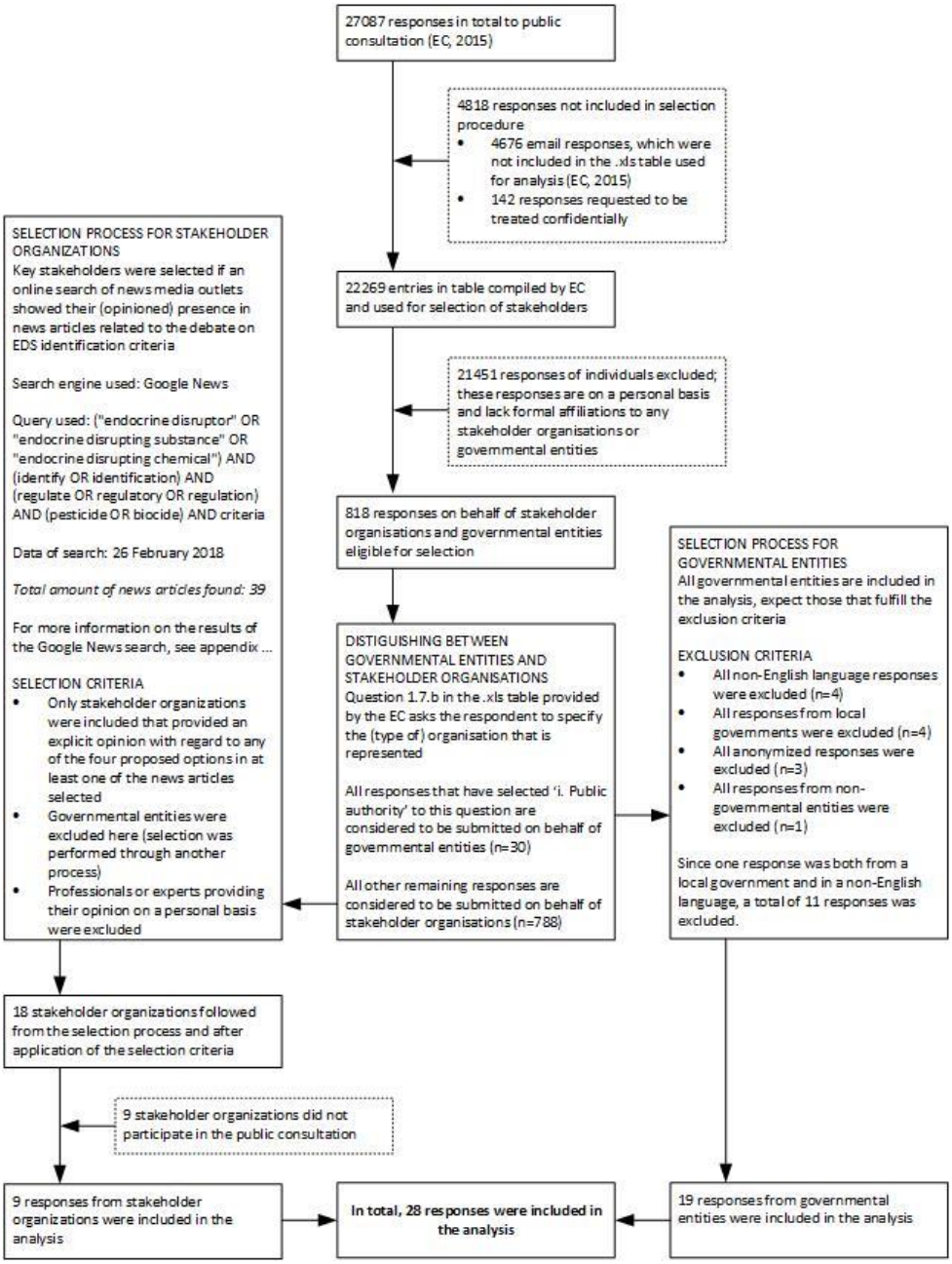

Figure 4.1: Flow chart of the procedure used to select the responses of stakeholder organizations and governmental entities.

other responses. The argumentation structures of all other responses were reconstructed by one author (LM) and reviewed by a second author (SCSC). Note that we only analyzed the contents of the responses as these were presented in the database of the Commission. No other documents attached to these responses were included in the analyses.

4.2.5 Identification of themes

Based on our earlier work, we are particularly interested in distinguishing arguments related to scientific knowledge from arguments related to evaluations based on normative values (see also Figure 2.12). In the context of this study, the starting point for the identification of *themes* was one *theme* touching upon the scientific understanding of EDSs, and one *theme* related to the more normative and political considerations surrounding the debate on the EDS identification criteria.

4.2.6 Categorization of arguments

After the identification of the *argumentation structures*, each individual argument in each of the *argumentation structures* received preliminary labels that were used to group them into 'preliminary argument categories'. That is, each argument was provisionally labelled to gain broad insight in the breadth and scope of all topics addressed. These provisional labels were then reviewed, and the wide variety of labels was reduced to a smaller amount of 'preliminary argument categories' by grouping similar labels under one category. Accordingly, all arguments in all argumentation structures were appointed to such an argument category. Note that these labels and categories were only used in this part of the analysis, as an intermediate step to later distinguish the ultimate *themes* and underlying *issues*. The argument categorization procedure is also shown in Figure 4.2.

4.2.7 Identification of underlying issues

The initial list of 'preliminary argument categories' appeared to be too crude. That is, most of these argument categories spanned multiple relevant topics that should be analyzed and discussed individually. For each argument, it was identified which *issue* it addressed. Through the process of filling the categorization table response-by-response, thereby first appointing each argument to one of the preliminary argument categories, and subsequently adding to each argument the *issue* it touched upon, the list of *themes* and *issues* was extended, revised and refined through an iterative process. This process was performed by one author (SCSC), with regular consultation of co-authors IvK, TGV, AHP and EL.

4.2.8 Identification of contrasting positions

During the categorization process, the contrasting *positions* towards the identified *themes* and *issues* became clear. For each *theme*, such *positions* were discerned. Each individual argument was then assigned to one of the two *positions* and coded with directional arrows and colors (see also document A2). This procedure was performed separately by two authors (SCSC and AHP), with regular consultation of co-authors IvK, TGV and EL.

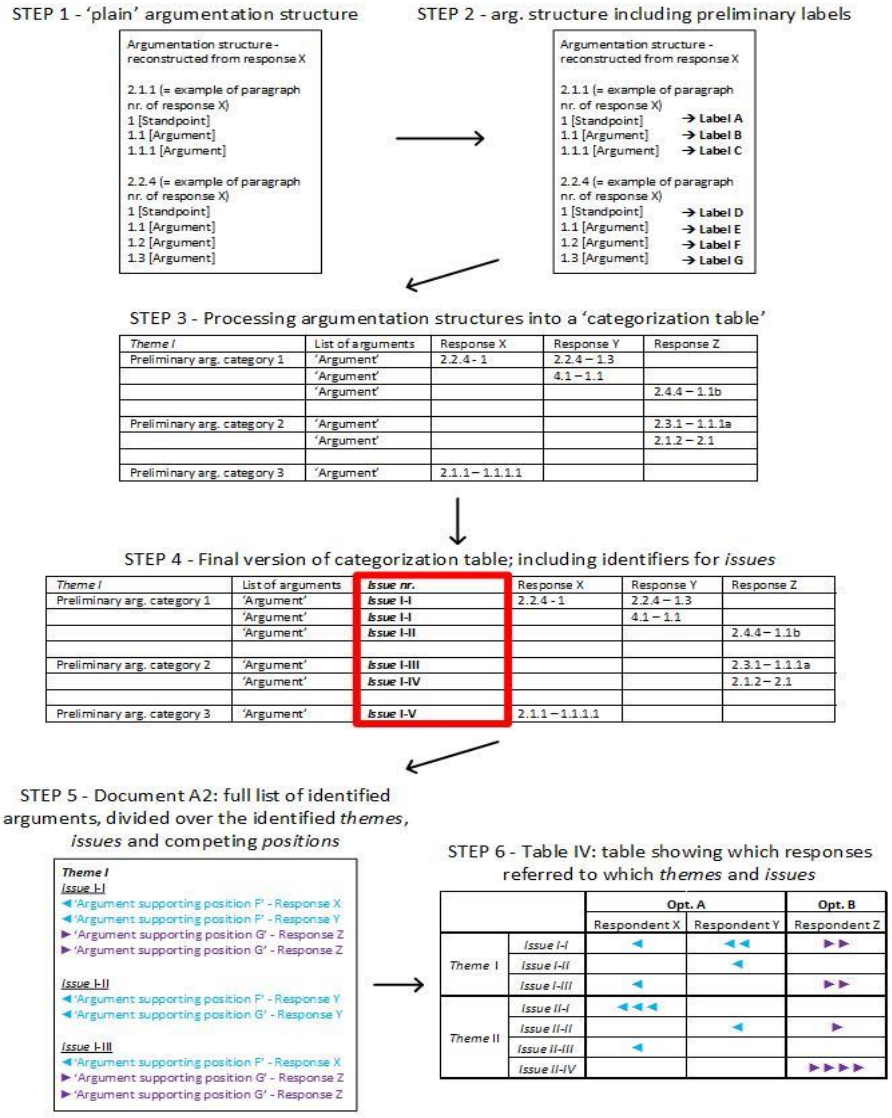

Figure 4.2: Simplified overview of the breakdown of argumentation structures into *themes*, underlying *issues* and particular *positions* of responses. Step 1 included the generation of all argumentation structures. To each of the standpoints and arguments in these argumentation structures, preliminary labels were added (step 2). In step 3, one document (categorization table) containing all arguments was developed, where all arguments were categorized into 'preliminary argument categories'. In step 4, *issues* were discerned in the categorization table to further distinguish relevant topics referred to in the responses. In step 5, the list of arguments presented in paragraph A4.2 in the appendix was developed, which contains a list of all *themes* and *issues*, and the arguments considered to address these *themes* and *issues* was developed, including the *position* that the argument refers to (indicated by different colors and arrows). In step 6, Table 4.4 was developed to summarize the information of the information in paragraph A4.2 of the appendix into a table.

4.3 RESULTS

4.3.1 Characterization of the responses

From the selection process, 28 responses were identified as being eligible for this study. Two EU member states had separate responses from different national agencies, while the standpoints and supporting argumentation put forward were essentially the same. The Danish VFA response refers to the contents of the Danish EPA response, so these responses were considered as one. Also, the responses of the two Austrian agencies were taken together, as these were practically identical. Accordingly, 26 unique responses were taken forward in the analyses; 17 of governmental organizations and 9 of stakeholder organizations (see Table 4.2).

Six types of option preferences were discerned from the identification of preferred policy preferences in each response. These all concern Options 3 and 4, or variations thereof, of the Roadmap of the Commission (EC, 2014); none of the included responses indicated a preference for Options 1 or 2. A brief discussion of the (predominantly negative) evaluation of Options 1 and 2 can be found in the appendix (see paragraph A4.3).

'Category approach' refers to those responses in which Option 3 of the Roadmap (i.e. adopting categories for different weights-of-evidence) is explicitly and unambiguously supported. 'Different variations of category approach' refers to responses that favor different altered versions of Option 3. 'Including potency' refers to responses that support Option 4 of the Roadmap (i.e. the inclusion of a potency consideration, in addition to the 2002 WHO/IPCS EDS definition). 'Including potency, and additional elements of hazard characterization' refers to responses that require Option 4 to be supplemented with additional elements of hazard characterization, such as severity and reversibility of effect; the inclusion of potency as the only element of hazard characterization is not considered sufficient to distinguish substances of high regulatory concern from those of low concern. 'Risk-based option' refers to an option not included in the EC's roadmap, but which is added here to reflect respondents' preference for risk-based, rather than hazard-based identification criteria. The essential difference with Option 4 of the Roadmap is the inclusion of exposure considerations. 'No specific preference' refers to those responses that either intentionally do not provide a specific option preference, or where a preference could not be reliably discerned from the response.

4.3.2 Identification of themes, underlying issues and contrasting positions

Five *themes* and 21 underlying *issues* were identified, as well as five sets of two contrasting *positions* that represent the two opposing perspectives as to each *theme* (see also Table 4.3).

Firstly, different *positions* were identified about whether the mechanistic activity and toxicological properties of EDSs are in fact different from those of other types of potentially hazardous substances, such that EDSs require specific study designs (*Theme 1*). We identified more different arguments supporting the *position* that EDSs have highly specific toxicological properties than arguments supporting the *position* that EDSs have toxicological properties that are not different from those of other potentially hazardous substances. Underlying *issues* are related to the timing of exposure and effects, dose-response relationships, mixture effects and

Table 4.2: Overview of the 26 responses included in this study. 'Preferred policy option' refers to one of the six identified option preferences. 'Respondent name' and 'respondent type' are as reported in the response. The label 'coordinated' was added when the response was submitted on behalf of an entire national government. The labels for 'respondent type' are based on the options provided in the consultation. Abbreviations: i. PA (EU)/(EEA)/(Non-EU) – Public authority located within the EU, in the European Economic Area or outside of the EU, respectively; vi. C/NGO – Consumer/Non-Governmental Organization; vii. I/TO - Industrial or trade organization.

Preferred policy option	Respondent name	Respondent type	ID Nr.
Category approach	French coordinated	i. PA (EU)	1
	Danish EPA/VFA	i. PA (EU)	2
	Finnish TUKES	i. PA (EU)	3
	Swedish KEMI	i. PA (EU)	4
	Norwegian FSA	i. PA (EEA)	5
	BEUC	vi. C/NGO	6
	HEAL	vi. C/NGO	7
	PAN Europe	vi. C/NGO	8
	Endocrine Society	viii. Other	9
	EurEau	viii. Other	10
Different variations of category approach	Belgian coordinated	i. PA (EU)	11
	Dutch coordinated	i. PA (EU)	12
	German UBA	i. PA (EU)	13
Including potency	German BfR	i. PA (EU)	14
Including potency, and additional elements of hazard characterization	UK coordinated	i. PA (EU)	15
	BCPC	vi. C/NGO	16
	CEFIC	vii. I/TO	17
	ECPA	vii. I/TO	18
	PlasticsEurope	vii. I/TO	19
	Australian coordinated	i. PA (non-EU)	20
	Canadian coordinated	i. PA (non-EU)	21
	New Zealand coordinated	i. PA (non-EU)	22
	US coordinated	i. PA (non-EU)	23
	Health Canada	i. PA (non-EU)	24
No specific preference	Austrian AGES/UBA	i. PA (EU)	25
	Hungarian NICS	i. PA (EU)	26

the assessment of environmental EDSs. Of the four *issues*, only 'dose-response' and 'environment' elicited contrasting perspectives. Overall, most arguments were related to 'dose-response', suggesting that this may be the most contested (see Table 4.4).

Secondly, we identified differences in *positions* as to the level of weight of evidence available and necessary to identify EDSs in accordance with the proposed regulatory options (*Theme 2*). Underlying *issues* were the availability of EDS-related scientific data, their quality and variability, the strength of evidence to establish causality, and the use of data on environmental EDSs. In support of the *position* that the identification of EDSs should occur on the basis of 'lower' weight of evidence requirements, arguments addressing the 'availability of EDS-related data' were used most. To support the *position* that the identification of EDSs should occur on the basis of a 'higher' weight of evidence requirement, arguments addressing considerations of causality establishment make up the majority.

Thirdly, we recognized different perspectives as to the consistency of a category approach or an option including potency with existing regulatory frameworks (*Theme 3*). Underlying *issues* are related to their application in the PPPR, BPR, CLP, REACH industrial chemicals regulation and Cosmetics regulations, and to the usefulness of distinguishing EDSs as a distinct substance category that requires specific regulatory attention. To support the *position* that a category approach is most consistent, or including potency is least consistent with existing regulatory frameworks, arguments addressing the consistency with PPP/BP regulations were used most. In support of the contrasting *position*, arguments addressing the usefulness of categorizing EDSs were used most.

Fourthly, we noted that specific properties of a category approach appeared to be evaluated differently by different respondents (*Theme 4*). Underlying *issues* dealt with the practicality of applying a weight-of-evidence-based category approach, the anticipated consequences to the identification of substances and expert judgment processes under this option, the anticipated impact on animal testing needs, and its applicability to environmental EDSs. Arguments addressing the practicability of a category approach were used most in support of the two contrasting *positions* (i.e. specific aspects of a category approach make this option most or least favorable, respectively).

Fifthly, specific properties of an option including potency also appeared to be evaluated differently by different respondents (*Theme 5*). Underlying *issues* are related to the practicality and consequences of applying a potency threshold approach, anticipated consequences of the inclusion of potency and its applicability for environmental EDSs. Arguments addressing the practicability of including potency were used most in support of the two contrasting *positions* (i.e. specific aspects related to including potency make this option most or least favorable, respectively).

Document A2 shows how all arguments are related to these *issues*, and in which *positions* this resulted. Table 4.4 summarizes this data. Contrasting *positions* were made visible by different

Table 4.3: Overview of five *themes*, 21 underlying *issues* and contrasting *positions*. Upper part: blue left- and purple right-pointing triangles indicate the major contrasting *positions*. Lower part: green up- and red down-pointing triangles indicate the (un)favorability of a category approach or including potency considerations, the key elements in Options 3 and 4 of the EC's roadmap, respectively.

1: Mechanistic understanding of EDSs	1.1: Timing of exposure and effects	◄ EDSs have highly specific toxicological properties ► EDSs have toxicological properties that are not different from those of other potentially hazardous substances
	1.2: Dose-response	
	1.3: Mixture effects	
	1.4: Assessment of environmental EDSs	
2: Regulatory considerations related to the idenfication of EDSs	2.1: Availability of EDS-related data	◄ Identification of EDSs should require a relatively lower weight of evidence ► Identification of EDSs should require a relatively higher weight of evidence
	2.2: Quality and variability of data on EDSs	
	2.3: Establishing causality	
	2.4: Assessment of environmental EDSs	
3: Consistency with existing regulatory frameworks	3.1: Consistency with PPP/BP regulations	◄ Category approach is most consistent, or including potency is least consistent with existing regulatory frameworks ► Including potency is most consistent, or a category approach is least consistent with existing regulatory frameworks
	3.2: Consistency with REACH regulation	
	3.3: Consistency with CLP Regulation	
	3.4: Consistency with Cosmetics Regulation	
	3.5: Usefulness of categorizing EDSs as a specific regulatory substance category	
4: Evaluations of specific issues related to a category approach	4.1: Practicality of applying a category approach	▲ Specific aspects of a category approach make this option most favorable ▼ Specific aspects of a category approach make this option least favorable
	4.2: Anticipated impact on the categorization of substances	
	4.3: Anticipated impact on expert judgment processes	
	4.4: Anticipated impact on the amount of animals used in animal testing	
	4.5: Suitability of a category approach for dealing with environmental EDSs	
5: Evaluations of specific issues related to including potency	5.1: Practicality of including potency	▲ Specific aspects related to including potency make this option most favorable ▼ Specific aspects related to including potency make this option least favorable
	5.2: Anticipated consequences of including potency	
	5.3: Suitability of including potency for dealing with environmental EDSs	

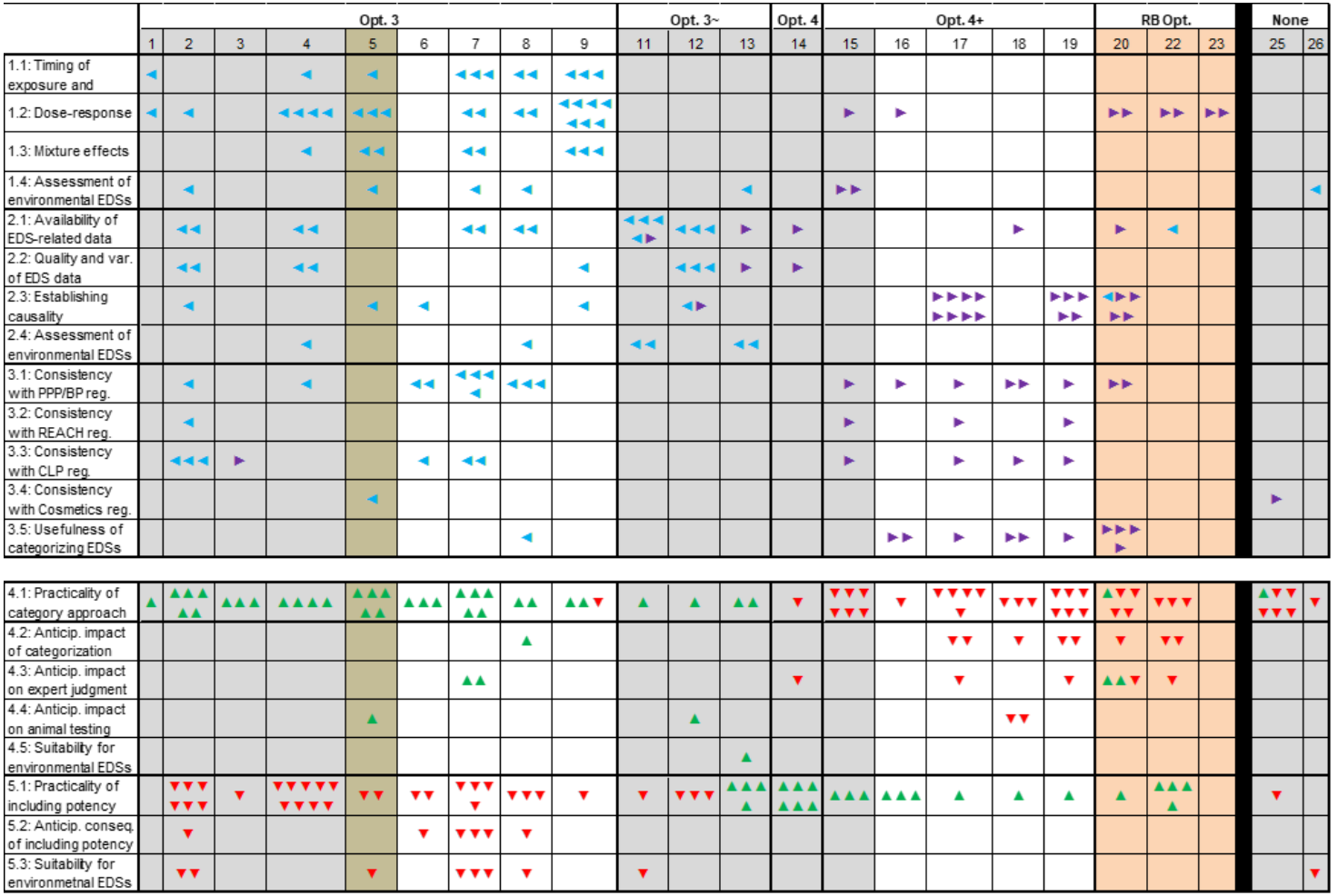
Opt. 3
Opt. 3~
Opt. 4
Opt. 4+
RB Opt.
None
1
2
3
4
5
6
7
8
9
11
12
13
14
15
16
17
18
19
20
22
23
25
26
1.1: Timing of exposure and
1.2: Dose-response
1.3: Mixture effects
1.4: Assessment of environmental EDSs
2.1: Availability of EDS-related data
2.2: Quality and var. of EDS data
2.3: Establishing causality
2.4: Assessment of environmental EDSs
3.1: Consistency with PPP/BP reg.
3.2: Consistency with REACH reg.
3.3: Consistency with CLP reg.
3.4: Consistency with Cosmetics reg.
3.5: Usefulness of categorizing EDSs
4.1: Practicality of category approach
4.2: Anticip. impact of categorization
4.3: Anticip. impact on expert judgment
4.4: Anticip. impact on animal testing
4.5: Suitability for environmental EDSs
5.1: Practicality of including potency
5.2: Anticip. conseq. of including potency
5.3: Suitability for environmetnal EDSs

Table 4.4 (previous page): Criteria options (first row) and respondents (second row; see Table 4.2 for the list of respondents) plotted against the list of *issues* (abbreviated representation in first column; see Table 4.3 for the list of *themes* and full name of the *issues*). The color of the column represents the type of respondent. All coloured columns are governmental entities: grey – EU member state, brown – member of the European Economic Union, orange – non-European government. Colorless columns are all stakeholder organizations. Number of triangles represent the number of arguments given per *issue* per response. Blue left- and purple right-pointing triangles (upper part), and green down- and red up-pointing triangles (lower part) indicate contrasting positions (see Table 4.3 for the list of contrasting *positions*). Columns of respondents 10, 21 and 24 were removed, since these would be empty (i.e. none of these responses included arguments related to the identified *issues*). Abbreviations: Opt. 3 – Category approach; Opt. 3~ - Different variations of category approach; Opt. 4 – including potency; Opt. 4+ - Including potency, and additional hazard characterization elements; RB Opt. – Risk-based option; None – No specific reference.

types of triangles. The amount of triangles shows how many arguments associated with an *issue* were found in each response. Note that the amount of triangles depicted for each response also depends on the length of that response. In addition, the representation in Tables 4.3 and 4.4 of *themes* 1, 2 and 3 is different from that of *themes* 4 and 5, since the first three *themes* relate to general scientific, regulatory scientific and regulatory aspects and the fourth and fifth relate to specific properties concerning the options.

4.4 DISCUSSION

We analyzed responses of 17 governmental entities and 9 stakeholder organizations to the Commission's public consultation related to the impact assessment of four options (proposed by the Commission) for criteria to identify EDSs for regulatory purposes. We used PDAT to identify the argumentation in support of the option preferences of the respondents. Through this analysis, we identified 21 *issues* that could be grouped into five *themes*. These five *themes* were: 1) the mechanistic understanding of EDSs, 2) regulatory considerations related to the identification of EDSs, 3) consistency of the options with existing regulatory frameworks, 4) evaluations of specific issues related to a category approach and 5) evaluations of specific issues related to the inclusion of potency.

4.4.1 Strengths and limitations

As far as we are aware, this is the first study that uses scientific argumentation analysis to better understand the variety of responses to the Commission's public consultation on impact of the EDS identification criteria. From the perspective of argumentation theory, the responses were not inherently argumentative by nature. Many responses consisted of a mix of (seemingly) non-argumentative information (e.g. information about the respondent), general comments towards or observations about options, and arguments put forward in favor or against a certain regulatory option. However, PDAT can only be used to study utterances that have an argumentative function (as compared to other communicative functions). Given the controversies and ensuing ambiguities surrounding EDS science and policy, we think it warranted to employ the strategy of *maximally argumentative interpretation* (van Eemeren, Grootendorst and Henkemans, 2010) in this study. This means that borderline cases that may or may not have been intended as argumentative were nevertheless considered as such in the analyses.

Although we have strived to maximize impartial and unbiased analysis and categorization of arguments, the interpretation and understanding of the subject matter by the researchers inherently may have had an influence. We are aware that the influence of the researchers' personal or professional biases can never be eliminated entirely in this type of analysis. However, several steps were taken to minimalize these influences: 1) The actual argumentation analysis was performed by researchers (LM and SCSC) who are not involved and have no position or stake in the ongoing research and policy initiatives regarding EDS. This analysis was supervised by an expert in PDAT (BG), 2) the PDAT method is geared towards performing analyses that remain true to the essence of the text that is analyzed, 3) judgments of the 'truth value' and 'comparative weights' of the (premises used in the) identified arguments were explicitly out of the scope of this article, due to the highly subjective nature of such evaluations, 4) the structured procedure for categorizing all arguments enabled us to retrace the steps followed and choices made during the categorization process and 4) almost all steps were performed by two or more authors. An exception is the development of the categorization table, but also here the findings were repeatedly corroborated with co-authors IvK, TGV, AHP and EL.

The Commission's consultation was performed to support the impact analysis of their four proposed regulatory options. Although the consultation did not focus on scientific and regulatory considerations of the options per se, our analysis shows that the responses put forward a wealth of arguments related to these topics. Accordingly, this consultation provides a rare opportunity to analyze and compare the argumentation and explicit regulatory preferences of a wide range of influential actors in the EDS science and policy debates. The Commission released a report on the results on the public consultation (see EC, 2015), with the aim to discuss the results of the consultation in the context of an impact assessment. This is different from the aims of this paper.

We assumed that the twenty six analyzed responses sufficiently represent the range of 818 responses on behalf of organizations. This point was most relevant for the selection process of stakeholder organizations, since all English language responses from non-anonymous, national governmental entities were included in this study. We consider that there is an inherently limited amount of influential stakeholder organizations (particularly NGO's, umbrella or lobby organizations active at the EU level) that will be both able and willing to dedicate the time, resources and expertise to have an active presence in the complex scientific, societal and regulatory debates on the EU's EDS regulation. We consider that online media presence, in terms of participation in newspaper articles or opinion papers appearing in the media, is an adequate proxy for this 'active presence' typical for influential stakeholders. In addition, we found that several non-included respondents explicitly referenced the responses of such key EU-level umbrella or lobby organizations as representing their official policy position. For example, we identified about 16 additional responses not included in our analysis (of both industrial parties and national-level trade unions) that refer to the CEFIC response as completely, or at least partly, representing their specific position.

4.4.2 Observations regarding the substance of the identified themes and issues

Of the five themes that were identified, only *themes* 1, 2 and 3 are discussed here, since these touch upon generic science and regulatory issues related to EDSs.

4.4.2.1 The science underlying science-based policy (theme 1)

The necessity of the criteria to be 'science-based' is explicitly emphasized in many responses and appears to attract wide consensus. However, the diversity in arguments used indicate that there are multiple interpretations of what this 'scientific basis' should be. Most notable is the substantial attention for the issue 'dose-response'. This issue resembles a known difference in perspectives occurring in the scientific debate on EDSs. For instance, it is disputed whether monotonic dose-response curves should remain the standard for assessing the toxicity of EDSs, in accordance with the centuries-old toxicological paradigm that dose is the key determinant of toxicity (contrast Vandenberg et al., 2012 with Autrup et al., 2015 and Beausoleil et al., 2016). The option to use non-linear dose-response models in risk assessment procedures was discussed in the context of the US EPA's proposed rulemaking (US EPA, 2018) and recent supplemental notice (US EPA, 2020) on transparency in regulatory science.

Notably, arguments used in support of positions in line with both a category approach and an option including potency contained references to scientific opinions of EFSA (2013) and JRC (2013), both established scientific bodies of the EU. However, different *positions* were supported by referring to the same document, possibly through selective referencing or interpretative ambiguity of the content in the report. For example, the ECPA response refers to the following conclusion of EFSA, to support a risk assessment approach (i.e. including hazard identification, hazard characterization and exposure) to the identification of EDSs: *"… endocrine disruptors can be … subject to risk assessment, where both hazard and exposure are considered in regulatory decision making. This is also the conclusion reached by the EFSA Scientific Committee in their Scientific Opinion published in March 2013 (lines 577-581 of ECPA response)"*. Alternatively, the HEAL response refers to another conclusion of the same Scientific Opinion, to support their critical position towards the inclusion of potency and further elements of hazard characterization: *"page 42-43 of EFSA – there is no scientific basis to include severity, irreversibility, critical effect or potency in the identification of EDCs (lines 385-386 of HEAL response)"*. The response of the Danish EPA refers to EFSA's Scientific Opinion to support their criticism about option 4 (including potency): "*Option 4 is not in line with recommendations from: … - The EFSA scientific committee, which in its scientific opinion discusses potency considerations as a part of the hazard characterization, not as part of the (hazard) identification of endocrine disruptors (EFSA, 2013) (lines 474, 490-492 of Danish EPA response)"*. In the responses of HEAL, PAN Europe and the Danish EPA, the JRC is referenced as remarking that the inclusion of potency in any EDS regulatory identification criteria lacks a scientific basis. These examples show how scientific opinions that are generally considered as authoritative and scientifically legitimate can be used to substantiate a particular policy preference, either by allowing for multiple interpretations of their scientific definitions and content, or by attributing different weights to the arguments put forward.

In most responses, it is explicitly supported that their preferred option is 'science-based', although the use of scientific arguments varies substantially. This may stem from uncertainty in the scientific evidence, leading to interpretative ambiguity. It may also be driven by selective use of science-based evidence, to serve the needs of the 'advocacy coalition' to support their normative regulatory preferences.

4.4.2.2 Weight-of-evidence considerations for identifying EDSs (theme 2)

The European Union chemical substances legislations, such as PPPR, BPR and REACH have provisions that require producers and downstream users to generate a minimum set of safety data for the respective substance. This data will be used to assess whether a substance fulfills the ultimately adopted criteria. Presumably, the starting point for these criteria will be the WHO/IPCS definition, which will subsequently require data that fulfills its key elements. The second *theme* relates to the weight of evidence that is considered both required and achievable to identify a substance as an EDS.

In support of the *position* that the identification of EDSs should occur on the basis of 'lower' weight of evidence requirements, arguments mostly touched upon the limited availability of EDS-related data. Particularly the Dutch and Belgian government responses use arguments referring to data-gaps in the regulatory assessment of EDSs. This finding is consistent with Dutch literature on this issue, where a discrepancy between the testing guidelines included in the EU's relevant regulatory frameworks and the data necessary to fulfill the WHO/IPCS definition is observed (see RIVM, 2016). Several arguments related to the contrasting *position* referred to the abundance or sufficiency, or positive developments regarding (the generation of) EDS-related data.

To support the *position* that the identification of EDSs should occur on the basis of a 'higher' weight of evidence requirement, arguments addressing causality considerations are predominantly used. These are particularly related to availability of alternative explanations for causality, and accordingly, the weight of evidence required to establish an EDS as such. First, a wide variety of potential influences are considered to potentially provide alternative explanations for relationships between exposure to a substance and the incidence of endocrine disruption related adverse health effects. Examples are the existence of other biochemical and physiological mechanisms besides endocrine disruption and the effects of physiological stresses and physical interactions (e.g. temperature). Second, arguments referred to the need to identify EDSs on the basis of certain weight-of-evidence requirements. For example, in the responses of CEFIC and PlasticsEurope, it was stated that EDSs should only be identified "*when there are clear adverse effects unambiguously caused by a well identified and empirically described mode of action" (CEFIC response: lines 289-291; PlasticsEurope response: lines 258-260)*. Alternatively, a contrasting *position* was supported by arguments related to the difficulty of establishing causality between exposure to EDSs and ED-mediated adverse health effects. For example, arguments refer to limitations in current regulatory testing strategies, the ubiquitous nature of exposure to EDSs and the multi-factorial nature of EDS-related adverse effects.

The observation that there are different perspectives as to the weight of evidence that is required and achievable for identifying EDSs is in line with insights from our earlier work. Based on an argumentation analysis of two pertinent publications in the field of EDSs science, Clahsen, van Klaveren et al. (2020) found that different perspectives among EDS experts occurred about the weight of evidence required to propose a certain policy measure. One side argued that objective methods to evaluate the weight of evidence exist and that Bradford Hill's *criteria* are an adequate starting point to establish causality in an unequivocal manner, whereas the other side argued that objective methods to evaluate the weight of evidence do not exist and Bradford Hill's *viewpoints* cannot be applied unequivocally. The researchers assert that this difference in perspectives was at least partly a manifestation of normative ambiguity, a term that refers to differences in (ethical) norms and values (see e.g. Renn, 2008). Accordingly, they conclude that addressing such normative elements in the debate on EDS science may benefit more from opening up the debate to interested and affected parties, than by performing more research.

Since the contrasting *positions* identified here are similarly at least partly normative in nature, normative judgments and not just purely scientific judgments need to be made, for the final selection of the preferred option. In this, the public consultation of the Commission is instrumental. This argumentation analysis shows that normative value judgments, which are often left implicit, can, and should be made explicit and distinguished from the purely interpretative judgment of the underlying science. Accordingly, instances of normative ambiguity could be addressed through broader stakeholder approaches, such as 'extended peer-community' approaches (see e.g. Ravetz, 1999) or participatory discourses (see e.g. IRGC, 2005; Renn, 2008).

4.4.2.3 Consistency of EDS identification criteria with existing regulations (theme 3)

Various respondents note that compatibility with existing regulatory frameworks is an important requirement for the ultimate identification criteria. The analysis of arguments addressing the consistency of EDS categorization with existing regulatory frameworks (*theme 3*) shows that there appear to be different perspectives towards the compatibility of either a category approach or an option including potency with existing formal provisions, and the wordings thereof, contained in the PPP, BP, REACH, CLP and Cosmetics regulations of the EU.

Supporters of a category approach mostly focused on the consistency of this option with the PPP and BP regulations. This may not be surprising, since the criteria to identify EDSs were specifically developed in the context of these regulatory frameworks. While supporters of an option including potency also addressed a range of arguments related to the consistency of the PPP and BP regulations with this option, most arguments addressed the usefulness of categorizing EDSs as a specific regulatory substance category. This is in line with the *position* referred to by several of these respondents in the context of *Theme 1*; from a toxicological point of view, EDSs are not necessarily different in terms of their toxicological properties, thus, it is not necessary to deal with EDSs differently from a regulatory perspective.

To support the *position* that either a category approach or an option including potency is most compatible with the PPP and BP regulations, different sections of the legal texts were cited. The key sections of these regulatory frameworks, and the different phrases, which relate to the circumstances under which potentially endocrine disrupting PPPs or BPs are approved, are outlined in Table 4.5. Some responses referred to the 'may cause' wording of the PPP and BP regulations. It is asserted that to identify an EDS in accordance with the WHO/IPCS definition, which includes the phrase 'consequently causes', requires a different and higher weight of evidence than the provisions of the PPP and BP regulations, which include 'may cause' wording. Adopting multiple categories that require lower weights of evidence than required for 'confirmed' EDSs is then considered much more appropriate, since this would enable the identification of 'potential' or 'suspected' EDSs, for which there may be indications of endocrine disrupting properties, but insufficient evidence to establish the substance as such. Alternatively, responses identified as favoring an inclusion of potency, inclusion of additional hazard characterization elements or the related risk- based option emphasized the 'negligible exposure' and 'negligible risk' exclusion clauses in the PPP and BP regulations, respectively. It was asserted that any assessment of EDSs under the PPP or the BP regulations would logically require a risk-based approach, since the exclusion clauses could inherently not be triggered without an assessment of relevant exposures. The use of references to specific provisions contained in existing regulatory frameworks was also identified in arguments about the consistency with other regulatory frameworks dealing with chemical substances (i.e. the REACH, CLP and Cosmetics regulations).

Table 4.5: The key sections and different key phrases of the PPPR and BPR that have been referred to by different respondents to argue for either a 'lower' or 'higher' weight of evidence approach to the identification of EDSs.

Legislation	Key section	Key phrases
PPPR	*"An active substance, safener or synergist shall only be approved if, … it is not considered to have endocrine disrupting properties that **may cause** adverse effect in humans, unless **the exposure** of humans to that active substance, safener or synergist in a plant protection product, under realistic proposed conditions of use, **is negligible**, …" (Regulation (EU) No 1107/2009, Annex II, 3.6.5).*	'may cause' adverse effects 'negligible exposure' as exclusion criterion
BPR	*… "active substances which, … are considered as having endocrine-disrupting properties that **may cause** adverse effects in humans … may be approved if it is shown that … **the risk** to humans, animals or the environment from exposure to the active substance in a biocidal product, under realistic worst case conditions of use, **is negligible**" (Regulation (EU) No 528/2012, Article 5).*	'may cause' adverse effects 'negligible risk' as exclusion criterion

4.4.3 Observations regarding the similarities and differences between the responses

We observed clear differences between the option preferences of different governmental entities, and a similarly clear division between the arguments used to support the different *positions* that are in line with these preferred options. In all responses of the governmental entities that have been identified as explicitly supporting a category approach, or variations thereof, (except that of the German UBA), arguments are used that consistently illuminate 'this side of the coin'. By contrast, all responses identified as supporting an option including potency, an option including additional hazard characterization elements or a risk-based option used arguments that consistently address 'this other side of the coin'. For *themes* 1, 2 and 3, virtually all *issues* addressed in support of a category approach (or variations thereof) had blue-left arrows, while simultaneously contra-arguments (purple right-arrows) were expressed in support of an option including potency, or the related two options (Table 4.4). Similarly, for *themes* 4 and 5, responses identified as being in favor of a category approach (or variations thereof, again excluding German UBA) consistently addressed arguments supporting this preference (green up-arrows), and against an inclusion of potency (red down-arrows). In responses identified as being in favor of an option including potency, a further supplemented hazard characterization option or a risk-based option, the opposite was found, where arguments in favor of including potency (green up-arrows), and against a category approach (red down-arrows) were consistently identified.

It appears that there is particular overlap between the arguments used in responses favoring a category approach. The similarities between the arguments put forward and *issues* addressed in, on the one hand, the French government and Nordic governmental institution responses and, on the other hand, the responses of the NGO's BEUC, HEAL and PAN Europe and the professional medical organization Endocrine Society are notable. Thus, one might argue, the category approach is supported by these parties for generally the same types of expressed reasons.

Where most national governments provided input to the consultation with 'one voice' (primarily through submitting a coordinated response, or a response from one key governmental institution), there are two complementary and partly divergent German responses. Both bring forward mostly arguments on *issues* related to considerations related to the identification of EDSs (*theme* 2) and evaluations of specific aspects related to including potency (*theme* 5). Accordingly, the German UBA and the German BfR both state that an option including potency is most suitable for managing EDSs that pose a risk to human health. Since the German BfR is mostly concerned with human health, their preference for an option including potency may not be surprising. However, the German UBA, also concerned with environmental stressors, notes that only a category approach is suitable for dealing with environmental EDSs. Accordingly, a category approach is the ultimately preferred option of the German UBA. Thus, it appears that the seemingly divergent *positions* are well in

line, when viewed from the perspective of the organizations' remit, where UBA primarily takes environmental stress into consideration, while BfR focuses on human health.

Although the Commission proposed only one option that included one additional hazard characterization element (i.e., an option including potency), two additional versions of this option were observed in the analyzed responses: an option including further additional hazard characterization considerations, and a risk-based option that includes exposure considerations. Interestingly, geographically-based differences can be observed between the proponents of the three versions. The hazard-based responses have all been submitted by EU-based actors, whereas the risk-based option responses have all been submitted by non-EU governmental entities. Notably, all non-EU governmental entity responses appeal to the risk-based regulation of EDSs in their own countries, and generally criticize a hazard-based approach as being unnecessarily restrictive. Alternatively, the response of ECPA (18) may be exemplary for actors that would ultimately prefer a risk-based approach to EDSs regulation in the EU, but for pragmatic reasons opt for an option including potency, and additional elements of hazard characterization, which comes closest to a risk-based option but is still essentially hazard-based in nature. "Current EDS regulation in the EU is hazard-based and a switch to risk-based regulation would require complex legislative changes" is one of the arguments brought forward in the ECPA response.

Our results point to the existence of two overarching 'advocacy coalitions', consisting of the respondents that either have a preference for a category approach (or variations thereof) or for an option including potency (or one of the related two options). Among these coalitions, a wide range of arguments are put forward to support contrasting *positions* concerning scientific, regulatory scientific and regulatory arguments (i.e. *themes* 1 to 3) to support their contrasting option preferences. It should be noted that these coalitions may not be entirely homogeneous. With regard to the respondents that are ultimately sympathetic to a category approach, the French and Scandinavian governments and all NGO's used slightly different types of arguments and supported this option more unequivocally than the Belgian and Dutch governments and the German UBA, that included some reservations and require some adjustments. In addition, There are some (geography-based) differences between the arguments put forward and the exact options supported by the German BfR (supporting an option including potency), the industrial trade organizations and the U.K. government (supporting an further supplemented hazard characterization option) and the non-EU trade partners (supporting a risk-based option not included in the consultation). Since this study using argumentation analysis is inherently cross-sectional in nature, it studies the state-of-affairs at one static moment in time and cannot assess any actual interaction between the respondents, this study cannot discern 'advocacy coalitions' as originally intended by Sabatier and Jenkins-Smith (1993). Future research should focus on the actual degree of interaction between responses considered part of the same 'advocacy coalition'.

The proposed existence of (implicit) advocacy coalitions, and the wide variety of disagreements among these coalitions, raises questions as to how proceed further with the issue of the EDS

identification criteria. It should be noted that substances associated with endocrine disrupting properties are widely used, and societal impacts associated with these uses may be significant in terms of adverse health effects and environmental stress, but also with regard to economic well-being, competitiveness of chemical industries and innovative potential. We made explicit how arguments related to scientific knowledge were used in conjunction with arguments related to normative value-judgments. On this basis, we argue that, for the ultimately adopted EDS identification criteria to become accepted in the EU society, the debate on these criteria would benefit from a societal dialogue. Here, the various scientific, regulatory scientific and regulatory aspects should be explicitly approached as interrelated themes. This dialogue should be open to all interested and affected parties, and could be performed in accordance with inclusive approaches as proposed in contemporary risk governance literature (see e.g. IRGC, 2005; Renn, 2008).

4.5 CONCLUSION

In an analysis of the EU's public consultation related to the impact assessment to select identification criteria for the regulation of EDSs, five topical *themes* and 21 underlying *issues* were identified. For each *theme*, two contrasting *positions* were discerned; one most in line with a preference for a category approach (or variations thereof), and one most in line with a preference for including potency (or related options). Accordingly, we argue that two overarching (implicit) 'advocacy coalitions' can be identified, using a wide range of contrasting arguments, related to the five identified *themes*, to support their preferred option. Among these 'coalitions', there appears to be consensus about the necessity of the ultimate option to be science-based, although different perspectives were identified as to what the most accurate mechanistic understanding of EDSs entails. We identified geography-based differences between the option preferences of countries; all responses of EU-based parties ultimately preferred hazard-based options, whereas the responses on non-EU-based parties preferred a risk-based option not included in the consultation. To move the discussion on EDS identification criteria forward, we argue that a societal dialogue would be beneficial, in accordance with contemporary risk governance literature, where EDS science and regulation are explicitly discussed as interrelated themes.

ACKNOWLEDGMENTS

The authors would like to thank Rik Bogers and three anonymous reviewers for their insightful comments on an earlier version of this article.

General discussion

5

5.1 INTRODUCTION

In this thesis, the issue of endocrine disrupting substances (EDS) was selected as research topic for two reasons. First, to perform an in-depth study of the reasoning and argumentation underlying contrasting regulatory preferences that appeared to exist among member states and stakeholders in the EU. Some of the parties involved envision strong regulatory action on EDSs and others prefer a more modest regulatory approach (see e.g. EC, 2015). Second, we wanted to study the role of the ongoing scientific controversy on EDSs in these contrasting regulatory preferences. The scientific literature remains largely undecided about what is the 'correct' interpretation of principles and concepts that are crucial for applying the scientific definition of EDS in practice, and subsequently, which substances can in fact be established as EDSs and be subjected to the applicable regulatory provisions. The central research question in this thesis was: *What argumentation underlies the international differences in preferences of stakeholders towards the regulation of EDSs?* This question was addressed in three separate studies; one extensive literature review covering a diverse and broad set of scientific fields and eight domain-specific conceptual frameworks, and two empirical studies. In the empirical studies, Pragma-Dialectical Argumentation Theory was used to perform argumentation analysis on the scientific controversy over EDS and the different stated regulatory preferences of government entities and stakeholder organizations.

5.2 OVERVIEW OF STUDIES AND MAIN RESEARCH FINDINGS

In **Chapter 2**, our theoretical review is presented. Eight conceptual frameworks were discussed that could provide eight different explanations for the occurrence of international differences in environmental health risk management. These were the Risk Assessment Paradigm, research into the roles of experts as policy advisors, the Psychometric Paradigm, the Cultural Theory of Risk, participatory approaches to risk assessment and risk management, the Advocacy Coalition Framework, the Social Amplification of Risk Framework, and Hofstede's Model of National Cultures. These conceptual frameworks developed mostly independently in different scientific fields, such as toxicology, epidemiology and risk assessment, the psychology or sociology of risk or policy sciences, over several decades, and most originating from the previous century.

The central aim of this review was to provide a theoretical foundation for empirical research that seeks to understand why countries frequently manage environmental health risks differently. Specific attention went out to study, from a theoretical perspective, how expert policy advisors specifically, and scientific knowledge in general, are involved in the development of these different risk management strategies.

From the ideas and concepts offered by the eight frameworks, pertinent questions were derived to be used in support of further empirical work (e.g. interviews or further argumentation analyses) focusing on international differences in environmental health risk management (see Table 2.2 in Chapter 2). The review also led to the development of an overarching framework that illustrates the conceptual relationships between the eight frameworks (see Figure 2.12). This overarching framework illustrates how competing *advocacy coalitions* can interact through

the exchange of arguments, and how these arguments are based on the professional and personal knowledge, beliefs and (cultural) contexts of the coalition's members.

In **Chapter 3**, a study is described where argumentation analysis was applied to two high-profile scientific papers taking opposing views on a 2012 update of the WHO report 'State of the Science on Endocrine Disrupting Substances'. The first was 'Critical comments on the WHO-UNEP State of the Science of Endocrine Disrupting Chemicals - 2012', by Lamb et al. (2014). The response to this paper was 'Manufacturing doubt about endocrine disruptor science – A rebuttal of industry-sponsored critical comments on the UNEP-WHO report 'State of the Science of Endocrine Disrupting Chemicals 2012', authored by Bergman et al. (2015). Both papers appeared in the journal Regulatory Toxicology and Pharmacology.

Our analysis began with describing how the discussion between the two author groups proceeded in argumentative terms, using Pragma-Dialectical Argumentation Theory (PDAT). Then, a range of *starting points* contained in both papers was identified. We defined the concept of *starting points* as a 'set of (typically unexpressed) knowledge, experiences, beliefs, norms and values that provides the basis for standpoints and argumentation put forward in a critical discussion', following the description of van Eemeren et al. (2010). We compared and contrasted these starting points across papers and identified whether these were cases of *interpretative ambiguity* about underlying scientific evidence or *normative ambiguity* about differences in broader norms and values. Two of the latter cases were studied in more detail using the typologies of Myths of Nature and Weiss' expert roles typology, to help make explicit the normative value differences as stake.

The analysis revealed five differences in starting points: 1) Whether the discussion falls into the category of a scientific controversy, 2) Whether objective 'weight-of-evidence' approaches exist, 3) Whether systematic, unequivocal approaches to establish causality exist, 4) Whether the endocrine system is a resilient system, designed to cope with environmental chemical exposures (within boundaries) or whether it is inherently vulnerable to such exposures, and 5) Whether state-of-the-science reports should be based on robust, objective and systematic scientific procedures, or whether these should be explicitly responsive to considerations of (precautionary) public health protection. All of these starting points, except perhaps for the first, included elements of *normative ambiguity*, meaning that differences in values and ethical norms were observed.

In **Chapter 4**, the second study employing argumentation analysis is described. Here, a range of responses to the EU's public consultation on identification criteria for endocrine disruptor regulation was analyzed using Pragma-Dialectical Argumentation Theory (PDAT). This study was triggered by the observation that, in different responses, different regulatory options were supported. The discussions mainly revolved around two options, that are both based on the WHO-IPCS definition of an EDS (i.e. *an endocrine disruptor is an exogenous substance or mixture that alters function(s) of the endocrine system and consequently causes adverse health effects in an intact organism, or its progeny, or (sub)populations*), but add different elements.

Some respondents prefer an option that adds considerations of potency, or even other elements of *hazard characterization*, whereas others prefer an option that avoids these considerations and introduces several weight-of-evidence categories. Accordingly, the main aim of this study was to analyze and compare the argumentation underlying different option preferences of governmental entities (e.g. national governments) and of prominent stakeholder organizations (e.g. NGO's or industry organizations), as stated in their responses.

Our selection procedure returned 26 responses for inclusion in this study (see Figure 4.2 and Table 4.2 in Chapter 4). We identified all individual arguments contained in these responses and subsequently categorized these arguments through an iterative process (see Figure 4.3). This ultimately yielded a list of five topical argumentation *themes* and 21 underlying argumentation *issues*. For each *theme*, two contrasting *positions* could be discerned (see Table 4.3 in Chapter 4); one most in line with a preference for a category approach (or variations thereof), and one most in line with a preference for including potency (or related options). For example, for *theme* 1), 'Mechanistic understanding of EDSs', we identified the *position* 'EDSs have highly specific toxicological properties', which was found predominantly in responses supporting a category approach, whereas the identified contrasting *position* 'EDSs have toxicological properties that are not different from those of other potentially hazardous substances' was particularly found in responses supporting the option including potency. The entire list of arguments underlying each of these *positions* can be found in Appendix A4.2. Besides 'Mechanistic understanding of EDSs', the other *themes* are 2) Regulatory considerations related to the identification of EDSs, 3) Consistency with existing regulatory frameworks, and 4) Evaluations of specific issues related to a category approach and 5) Evaluations of specific issues related to including potency.

The main finding is that two overarching (implicit) 'advocacy coalitions' could be discerned, that consistently appear to use arguments that support positions in favor of either the category approach option or an option including potency considerations. Among these 'coalitions', there appears to be consensus about the necessity of the ultimate option to be science-based, although different perspectives were identified as to what the most accurate mechanistic understanding of EDSs entails. Geography-based differences between the option preferences of countries were also identified; all responses of EU-based parties ultimately preferred hazard-based options, whereas the responses on non-EU-based parties preferred a risk-based option not included in the consultation.

5.3 WHAT WE RECOGNIZE FROM THE THEORETICAL REVIEW AND OTHER LITERATURE

Several conceptual frameworks discussed in the theoretical review, and some additional literature, can be used to reflect on our empirical findings. From these insights, some suggestions were distilled for a way forward in the debates on EDS science, policy and regulation.

5.3.1 Risk Assessment Paradigm: EDS and hazard (versus hazard and potency) versus risk

Substantial debate remains to be focused on whether EDS should be regulated on the basis of principles of hazard identification (HI), hazard characterization (HC) or risk assessment (RA). Briefly, HI is concerned with identifying the type and nature of the adverse effects caused by exposure to an EDS, whereas HC includes qualitative or quantitative measures of the potential adverse effects by deriving e.g. dose-response relationships or bench-mark doses (FAO, 2020). Where HC is still strictly concerned with the toxicological properties of the substance itself, RA also includes considerations of exposure, such as exposure routes and levels of exposure. It should be noted that the commonly accepted definition of an EDS - an endocrine disruptor is an exogenous substance or mixture that alters function(s) of the endocrine system and consequently causes adverse health effects in an intact organism, or its progeny, or (sub)populations – is HI by nature.

In Europe, current regulatory provisions of EDSs are hazard-based, and discussions about the identification criteria that the EC was establishing for EDS regulation were geared towards HI-based versus HC-based regulation. In fact, of the four options proposed by the EC, none of these were risk-based, that is, included exposure considerations. The reasoning is that when a substance is established as an EDS following the above HI-based definition, then exposure is essentially irrelevant; the potential impacts of EDS-type adverse effects are considered so detrimental, that no exposure level whatsoever is considered acceptable from a public health standpoint. Instead, as shown in Chapter 4, the discussions among EU governmental entities and stakeholder organizations centered on the concept of potency, a measure for toxic capacity, understood as the amount of substance necessary to induce an adverse effect of a given intensity. Potency could be considered a proxy for dose-response curves, and featured in the HC-type option proposed by the EC.

In countries outside Europe, mostly Australia, Canada, New Zealand and the U.S., risk-based regulatory approaches to EDSs were advocated in responses to the public consultation held by the EC. A general theme was that hazard-based regulations would be unnecessarily restrictive to trade. For example, in the response on behalf of the Australian government, it was argued that endocrine disruption does not constitute a specific end-point per se, but rather includes a range of mechanisms of action potentially inducing adverse effects that, in turn, are already routinely and effectively considered in the countries' existing – risk-based – regulatory provisions. Accordingly, the adoption of hazard-based criteria in the EU could have restrictive and therefore detrimental impacts on trade with countries having risk-based regulation for EDSs.

Hofstede's Dimensions of National Cultures could be used to study and explain these observed international differences in hazard versus risk-based regulatory approaches to EDSs. The argumentation analysis in Chapter 4 did not address these dimensions, but future research could study if the dimensions of power distance, uncertainty avoidance, individualism,

masculinity, long-term orientation or indulgence could (in conjunction) provide fruitful explanations.

5.3.2 Expert roles typologies and the Myths of Nature

In Chapter 3, the roles of two EDS expert groups and particularly any interpretative and normative ambiguity occurring between these groups were analyzed. With regard to a characterization of roles in terms of Pielke's typology, this is difficult, since both expert groups do not explicitly seek to explicitly provide policy advice, which Pielke's roles are mostly about. Rather, the discussion between the groups focusses on several key scientific and regulatory scientific issues on EDS. However, the papers do include some resolute phrases, such as 'this calls into question the integrity of decisions at all levels of the 2012 report' in Lamb et al. (2014) and 'manufacturing doubt about endocrine disruptor science' as part of the title of Bergman et al. (2015). Also, both expert groups explicitly adopt vocal positions towards, for example, the objectivity of methods to evaluate the weight of evidence, and the adequacy of using Bradford Hill's considerations for establishing causality. One could argue that the two expert groups would then most resemble 'issue advocates'.

With regard to Weiss' typology, similarities to the 'scientific absolutism' and 'environmental absolutism' expert roles were found. These expert roles refer to values about the Precautionary Principle and the degree of certainty required to advise a policy measure of a given stringency. We considered that the emphasis of Lamb et al. on robust, objective and systematic procedures for a state-of the-science report on EDSs, including supporting argumentation, shows elements of 'scientific absolutism'. Conversely, Bergman et al.'s emphasis on the explicit responsiveness to considerations of public health protection in a state-of-the-science report shows elements of 'environmental absolutism'. However, as noted in Chapter 3, whether the expert role typologies of Weiss (2003) do in fact accurately capture the identified difference in starting points needs further research, since both of the analyzed papers make little explicit references to preferred policy measures regarding particular EDSs.

In Chapter 3, we also used the Myth of Nature typology developed by Schwartz and Thompson (1990) to reflect on the two author groups' perspectives towards the resilience of the endocrine system. This typology includes different sets of values towards ecosystem stability. In short, we observed that Lamb et al.'s illustration of the endocrine system in human physiology resembles the 'nature is tolerant, but within limits' Myth of Nature. Arguments identified in Lamb et al. (2014) point to the understanding that a key function of the endocrine system is to deal with continuous fluctuations in hormone levels, though the limits of these homeostatic processes need to be respected. This perspective is much similar to the perspective of nature as a tolerant system that has its boundaries. Alternatively, arguments identified in Bergman et al. (2015) point to the illustration of the endocrine system as resembling the 'nature ephemeral or fragile' Myth of Nature. Bergman et al. emphasize the susceptibility of vulnerable groups to irreversible disruption of the endocrine system leading to potentially latent effects, even in situations of (very) low exposures to EDS.

5.3.3 Social Amplification of Risk Framework and social amplification stations

In Chapters 3 and 4, we did not specifically study whether social amplification or attenuation or rippling effects occurred in the debates on EDS-related substances and what effects these might have had on the general public. It was anticipated to apply the theoretical insights from SARF to the results of our research into the restriction proposal of BPA in thermal papers in France, but this line of research was not pursued further after gathering preliminary results (see Paragraph 1.5.2 in Chapter 1 for more information).

A range of stakeholder organizations could be identified as 'social amplification stations' that generate and transmit EDS related information through a range of communication channels. In Chapter 4, we performed an online search of news media outlets using Google News, to find the stakeholder organizations that had an 'opiniated presence' in the debates on the EU's identification criteria for EDSs, i.e., which were the most vocal about this topic in online media. From the list of organizations gathers from this media analysis, nine organizations in fact participated in the public consultation: three EU-level NGO's (BEUC, HEAL and PAN Europe), three EU-level industrial trade organizations (CEFIC, ECPA and PlasticsEurope), one British trade organization (BCPC), the global scientific community 'Endocrine Society' and European Federation of National Associations of Water Services 'EurEau'. Interestingly, seven of these organizations are either NGO's involved in the protection of human health and the environment or industrial trade organizations that aim to protect economic, employment and international trade concerns. This signals the significant divide in interests in the debate on EDSs regulation.

5.3.4 Advocacy Coalitions Framework: can we discern 'advocacy coalitions'?

In the theoretical review, we developed an overarching framework that visualizes the conceptual relationships between the included frameworks. This overarching framework (figure 2.12) consists of two levels. The upper level of the framework, dealing with the polemic occurring *between* advocacy coalitions, has received limited attention in this thesis. The focus of our study was on the lower level of the framework. We performed two empirical studies with a *cross-sectional* character, rather than a more *longitudinal* study in which the development of and the interaction between competing 'advocacy coalitions' is followed over time. Most interest went out to in-depth analyses of the argumentation put forward in the context of EDS science, policy and regulation. Since scientific and policy documents were used for the analysis of argumentation, and these documents portray the state-of-affairs at the time of writing, these analyses are inherently *cross-sectional* by nature. *Longitudinal* studies would require different types of research methods, such as interviews or focus groups. As discussed in Chapter 1, some interviews were performed to specifically address elements of the upper level of Figure 2.12. We did not pursue these elements further given our focus on the lower level of the framework.

The concept of *advocacy coalitions* plays a central role in our overarching framework. The concept was defined by Sabatier and Jenkins-Smith (1993) as '*people from a variety of positions (elected and agency officials, interest group leaders, researchers) who share a particular belief system - i.e. a set of basic values, causal assumptions, and problem perceptions - and who show a*

non-trivial degree of co-ordinated activity over time' (p. 25). From this definition, it becomes clear that an advocacy coalition can be considered as such when there is a non-trivial degree of coordinated activity, and when the members share a particular belief system. Then, the question arises whether the two author groups, whose work is subjected to argumentation analyses in Chapter 3 (i.e. Lamb et al., Bergman et al.) can be considered advocacy coalitions in accordance with the definition, and whether any advocacy coalitions can be discerned in the study outlined in Chapter 4.

It should be noted that argumentation analyses are not particularly suitable to discern a degree of coordinated activity, unless such activity is mentioned by the authors in their texts. However, one could argue that the act of drafting an elaborate scientific publication constitutes in itself a non-trivial degree of coordinated activity, given the substantial time that goes into the preparation, cooperation and coordination of such a manuscript, indicating a '*non-trivial degree of activity*'. It is more difficult to establish that this activity occurred '*over time*', considering that our study analyzed two documents appearing at one point in time. Notably, the Lamb et al. and Bergman et al. author groups are subgroups of bigger expert groups that have published scientific publications, in different compositions and during other periods of time, that appear to include similar (contrasting) viewpoints and reasoning as observed in the analyzed publications (see also paragraph 1.2 for a brief overview of key scientific issues). Examples are Rhomberg et al., 2012 vs. Kortenkamp et al., 2012, Dietrich et al., 2013 vs. Bergman et al., 2013 and Zoeller et al., 2014 vs. Autrup et al., 2015, which indicate at least some degree of non-trivial, and repeated, cooperation over time. Considering some values were made explicit (i.e., aspects of the belief systems) held by the Lamb et al. and Bergman et al. author groups, by using the Myth of Nature and Weiss typologies, these author groups could be considered *'de facto' advocacy coalitions* consisting primarily of scientific experts. Since the analysis was limited to the work of two author groups, we did not attempt to identify other, possibly less vocal expert groups that could have also formed *advocacy coalitions* through non-trivial degrees of coordinated activity, but which may hold different values or belief systems. This could be a subject for future research.

In Chapter 4, signs for the existence of *advocacy coalitions,* were also found as (at least) two broad collections of stakeholder organizations and governmental entities were identified. These put forward an apparently coherent line of contrasting positions, along with supporting arguments, concerning five topical themes. Though the belief or values systems underlying these positions and arguments were not made explicit, it could be argued that the observed differences in argumentation would not exist without different belief or values preferences. This premise is also made visible in Figure 2.12 in Chapter 2, which posits that argumentation referring to scientific knowledge, or argumentation referring to normative positions are inherently based on, among others, personal and professional beliefs that are invoked to interpret the situation. Therefore, it could be concluded that the contrasting positions and arguments identified are strong indications of the existence of contrasting *policy belief systems*. Since PDAT-based argumentation analyses were also used here, the criterion '*non-trivial coordinated activity over time*' is difficult to fulfil, though some types of coordinated activity

should have occurred between CEFIC (European Chemical Industry Council) and a range of individual chemical companies. These companies explicitly refer to the contents of the CEFIC response as completely, or at least partially, representing their position and argumentation. The same mechanism was observed with the response of ECPA (European Crop Protection Association), which was referred as conveying the position of several national and pan-national trade unions of pesticides producers. As with collectively drafting a scientific paper, it is expected that it took a significant amount of time, coordination and collaboration to prepare the elaborate CEFIC and ECPA responses, particularly to uphold their functions as representing a wide range of smaller and bigger companies that have substantial interest in adequate regulation of EDSs in Europe. It should be noted that we have little direct evidence of this cooperation on this topic occurring *over time*, though it is unlikely that the observed coordination efforts were performed just for this consultation; the 2014-2015 EU public consultation is one of many EU activities employed in the previous decades dealing with EDS science and regulation. Other evidence of governmental entities or different stakeholder organizations employing coordinated activities among one another was not found, and this could be an interesting direction for future research. All in all, the conclusion is that there are indications of '*de facto' advocacy coalitions*, though it appeared difficult to fulfill all conditions as put forward by Sabatier and Jenkins-Smith (1993).

5.3.5 Suggestions on how to proceed in debates on EDS science, policy and regulation

The present research project was initiated by experts of RIVM that, through their work as policy advisors to environmental and public health policy makers, made two observations that required theoretical and empirical study: 1) EDS science and policy appear to be intertwined through complex relationships, rather than being two separate realms as suggested by the traditional Risk Assessment Paradigm, and 2) there appear to be conflicts over normative values present in the scientific discourse on EDSs, and resolving these conflicts requires different means than the traditional remedy of performing additional research to bridge academic disputes. The current thesis largely corroborates these observations. However, this begs the question: what should a fruitful and constructive role of science be in establishing how EDSs should be regulated? Given the complexity of the answer to this question, this paragraph provides some suggestions.

First, it is important to consider that, in the context of highly complex environmental health risk issues like EDSs, scientific uncertainty is often misunderstood as an apparent lack of scientific understanding. For instance, it may be more fruitful to understand scientific uncertainty as a lack of consensus over the different existing scientific understandings, which, in turn, explicitly occur in a complex context of political stakes and a diverse landscape of cultural norms and values (see also Sarewitz, 2000). In this regard, the scientific debate on EDSs is archetypical; one only needs to use the search term 'endocrine disruption' in a scientific search engine like Scopus to yield thousands and thousands of publications somehow studying these substances. Following the often espoused belief that performing more research reduces

scientific uncertainty, one would expect that the uncertainty on the risks of EDSs has greatly diminished over time, yet the significant disagreements among experts persist.

On the part of the scientific realm, various scholars argue that scientific experts should be more transparent about their personal values and the values in their work. The key idea is that, since value influences are sometimes unavoidable, incorporating them explicitly and with scrutiny is a better alternative than leaving values implicitly present (Elliott, 2017). Transparency about values would also impact the policy advisory roles chosen by experts, since the consideration of roles essentially amounts to a subset of values specifically related to how a scientific experts considers his expertise would benefit himself or his surroundings. Pielke (2007) argues that all four ideal-typical expert roles are legitimate in scientific debates, under certain conditions, but that issue advocacy should not take the form of *stealth issue advocacy*, where an expert is advancing scientific argumentation, thereby using the authority of science to fulfill hidden values or (political) interests. By all accounts, scholars from a wide range of disciplines agree that dealing with the presence of values in science, particularly when these have already given rise to conflicts over norms and values among scientists, exceeds the 'jurisdiction' of scientific experts and requires broader stakeholder involvement and promoting scientific diversity and interdisciplinarity (Elliott, 2017; Hisschemöller and Hoppe, 1995; IRGC, 2005; Pielke, 2007; Ravetz, 1999; Renn, 2008; Sarewitz, 2004). For example, Ravetz (1999) suggests to establish 'extended peer communities', allowing all those interested and affected to be able to provide their critical evaluations in open dialogue settings, while Elliott (2017) refers to, among others, citizens serving as advisors in committees, community-based participatory research or incorporating social scientists and humanists in lab environments.

On the part of the policy realm, it could be argued that disputes over norms and values underlying environmental controversies should be settled in the political arena. The EU consultation on EDS identification criteria is a prime example of such an arena. Sarewitz (2004) asserts that politicians have persuasion, (re)framing, negotiation, rhetoric and other means at their disposal, and that scientific knowledge should be considered one of the many considerations on which policy makers can draw to make their decisions. Other valuable considerations could be political or economic interests, legal, religious, cultural or ethical considerations or personal worldviews, to name some. The responsibility of making decisions will then fully be with politicians, rather than scientific experts, whose scientific knowledge can in fact help to signal societal problems, formulate options for remediation or monitor the adequacy of the chosen policy measure. Various frameworks have been developed to provide guidelines for constructive interactions between scientific experts and policy makers. For example, the Risk Governance Framework of the IRGC (2005) outlines a structured, participatory governance to deal with complex, uncertain and ambiguous risk issues, where risk communication and stakeholder involvement are at the heart of the process. The NRC (1996) proposes a analytic-deliberative process of risk assessment and risk management, where the state-of-the-science is merged with continuous deliberation with interested and affected parties. Paragraph 3.2.5 of discusses these frameworks in more detail.

5.4 EXPERIENCES WITH PRAGMA-DIALECTICAL ARGUMENTATION THEORY

In the two empirical studies, pragma-dialectical argumentation theory (PDAT) was used to subject the selected documents to argumentation analysis. In each of these studies, we used the PDAT methodology to identify standpoints and underlying argumentation. The main products of the PDAT-based analyses were the identification of key standpoints and underlying argumentation structures of the documents under scrutiny. These findings were then used to make explicit the implicit interpretive and normative elements in the analyzed discussion (Chapter 3) or to be able to identify topical, broad themes and specific issues in a wide range of responses (Chapter 4).

Some strengths and some points for attention regarding the application of PDAT in this research project were encountered. An in depth discussion on PDAT, however, benefits from providing some additional technical information on its *pragmatic* and *dialectical* dimension, the pivotal 'ideal model for a critical discussion', the ten discussion rules of PDAT and the resulting options for *descriptive* and *normative* argumentative evaluations. After that, some reflections will be shared concerning our application of PDAT, what was learned from applying PDAT to the EDS case study, and from having life-sciences experts of RIVM collaborate with argumentation theory experts from the University of Amsterdam.

5.4.1 Pragmatical and dialectical dimensions of PDAT

As becomes apparent from its name, PDAT integrates from two lines of research orientations: *pragmatics* and *dialectics*. Scholars in the field of *pragmatics* study how 'real life' communication and interactions attain meaning within a certain context (van Eemeren et al., 2014). For example, the phrase 'I can't get this computer working' could be interpreted in multiple ways. The speaker could simply assert that he cannot get the computer to work. Alternatively, the speaker could also implicitly request help in getting the computer to work. The exact intended meaning of the phrase may then depend on the specific situation, identity of the speaker, intonation, body language, prior conversations and other contextual factors. PDAT's *dialectical* orientation refers to solving the difference of opinion in a *reasonable* manner (van Eemeren et al., 2014). In terms of the PDAT, *reasonableness* refers to committing sound argumentative moves that do not impede the resolving of a difference of opinion. Here, fallacies are understood as incorrect moves towards successfully resolving the dispute at hand (van Eemeren and Grootendorst, 1987).

5.4.2 Ideal model of critical discussion

The ideal model of a critical discussion of PDAT involves reconstruction of a discussion according to PDT's four stages of a critical discussion (van Eemeren, Grootendorst and Henkemans, 2010). This allows for an analysis of the complete argumentative interaction, instead of the analysis of single pieces of argumentation, as would be the case when using e.g. Toulmin's model of argumentation (Toulmin, 1958). The key implication of the ideal model is that discussants should go through all stages to maximize chances of reasonably solving a difference of opinion. These stages are the confrontation stage, opening stage, argumentative

stage and concluding stage (see table 5.1). The key implication of the ideal model is that discussants should go through all stages to maximize chances of reasonably solving a difference of opinion. If discussants miss one or more stages, these chances may drop considerably. For example, when discussants do not go through the confrontation stage, they may not have agreement on the exact standpoint that should be under discussion. Then, a Babylonian confusion of tongues is lurking, making it practically impossible to resolve the intended difference of opinion satisfactorily.

Table 5.1: The four stages of an ideal critical discussion

Confrontation stage	A difference of opinion is established, that is, one party's standpoint has been subjected to criticism or doubt by another party.
Opening stage	The parties engaged in the difference of opinion decide to resolve the conflict. They assign various roles (i.e., protagonist or antagonist) to one another and they establish whether there is enough common ground to build a discussion on
Argumentative stage	The protagonist(s) defend(s) his standpoint by putting forward argumentation and attempts to remove (anticipated) doubt and refute (possible) criticism from the opposing party.
Concluding stage	The extent of conflict resolution will be assessed, based on the acceptability of the argumentation. Ideally, the antagonist retracts the opposite standpoint (or doubt) if the protagonist has defended his standpoint successfully, or the protagonist retracts his standpoint if he fails to remove objections or doubt from the antagonist.

5.4.3 Ten discussion rules of PDAT

According to PDAT, and in accordance with its *dialectical dimension*, a critical discussion should proceed in accordance with ten discussion rules, in order to ultimately resolve a difference of opinion in a *reasonable* fashion (van Eemeren, Grootendorst and Henkemans, 2010), see also table 5.2. The key implication of these discussion rules is that violation of one of these rules will obstruct the path to reasonably resolving a difference of opinion. This will lead to fallacious reasoning. Well-known examples of fallacies are the *argumentum ad hominem* and the *argumentum ad ignorantiam*. In the former, the freedom rule is violated by attacking a discussant's personal credibility, rather than the contents of standpoints or arguments put forward by the discussant. The latter fallacy involves a violation of the closure rule. In common language, the *argumentum ad ignorantiam* is frequently described as 'absence of proof is not proof of absence' and features a false dichotomy. That is, some statement that has not been shown to be false, must be true (or vice versa), although 'inconclusive' or 'unresolved' may have been legitimate third options, but have remained implicit.

Table 5.2: The ten discussion rules of PDAT

1.	Freedom rule	Parties must not prevent each other from putting forward standpoints or casting doubt on standpoints
2.	Burdon of proof rule	A party who puts forward a standpoint is obliged to defend it if asked to do so
3.	Standpoint rule	A party's attack on a standpoint must relate to the standpoint that has indeed been advanced by the other party
4.	Relevance rule	A party may defend his or her standpoint only by advancing argumentation related to that standpoint
5.	Unexpressed premise rule	A party may not falsely present something as a premise that has been left unexpressed by the other party or deny a premise that he or she has left implicit
6.	Starting point rule	No party may falsely present a premise as an accepted starting point, or deny a premise representing an accepted starting point
7.	Argument scheme rule	A standpoint may not be regarded as conclusively defended if the defense does not take place by means of an appropriate argument scheme that is correctly applied
8.	Validity rule	The reasoning in the argumentation must be logically valid or must be capable of being made valid by making explicit one or more unexpressed premises
9.	Closure rule	A failed defense of a standpoint must result in the protagonist retracting the standpoint, and a successful defense of a standpoint must result in the antagonist retracting his or her doubts
10.	Usage rule	Parties must not use any formulations that are insufficiently clear or confusingly ambiguous, and they must interpret the formulations of the other party as carefully and accurately as possible

5.4.4 Descriptive and normative evaluations in the context of PDAT

Argumentative analysis in the context of PDAT can include both a *descriptive* and a *normative* dimension (van Eemeren et al., 2014); the argumentation is either 'plainly' described as interpreted by the analyst, or the degree of *reasonableness* of the argumentation is assessed, which requires to analyst to make some normative value judgments. *Descriptive* argumentative analysis involves reconstruction of a discussion in accordance with PDAT's four stages of a critical discussion and the identification of standpoints and argumentation and making unexpressed premises explicit is performed. A crucial step here, which has been described in Chapter 3, involves reconstructing the external and internal organization of the argumentation, referred to as the *argumentation structure* and *argument schemes*, respectively (van Eemeren et al., 2014). *Normative* evaluations are concerned with whether the discussion has indeed proceeded in a *reasonable* manner and evaluates to what extent the discussion indeed has resembled the ideal of a critical discussion. This involves an evaluation of the appropriateness of an argument scheme by asking *critical questions*, or specific evaluative questions that qualify the *reasonableness* for transferring acceptance from a premise to a standpoint (van Eemeren et al., 2014). Also the extent of violation of the ten rules of a critical discussion will need to be considered (van Eemeren, Grootendorst and Henkemans, 2010).

5.4.5 Reflections on how PDAT was (and was not) applied in this thesis

An important decision was that we would only focus on *descriptive* analyses of the texts under scrutiny, and refrain from performing *normative* evaluations. Our research aims, mostly concerned with painting an overall picture on the competing perspectives in the ongoing scientific and regulatory discussion, would most benefit from analyses that are performed as unbiased and impartial as reasonably possible. Normative evaluations about, for example, the 'truth value' of the technical substance of certain standpoints or arguments, or whether certain discussion rules were violated and fallacies were committed, would require us to make judgments, which would compromise our neutral position and may potentially influence our credibility as (argumentation) analysts, particularly in already high-stakes and polarized debates as those on EDS science, policy and regulation. This also means that we were unable to thoroughly include the *dialectical* dimension of PDAT, as we explicitly refrained from making evaluations of *reasonableness*. However, future research could explicitly focus on whether all discussion rules are obeyed in these debates on EDSs, and if not, which fallacies occur that could impede a reasonable solution to the pending differences of opinion. Even more so, fallacies could be committed purposefully as rhetorical means to persuade audiences about certain standpoints, without engaging in the technical substance of the issue. Particularly discussions about conflicts of interests of experts, which not seldomly feature in discussions related to EDS science and societal challenges like climate change and the potential risks of electromagnetic fields, could benefit from including analyses into fallacies and fallacious reasoning. It is key to differentiate between actual conflicts of interest that could severely impact the credibility of science and cases of *argumentum ad hominem* that could significantly undermine the credibility of the targeted expert.

By not performing *normative* evaluations, we did not assess to what extent the discussion observed in Chapter 3 fulfilled the four stages of an ideal critical discussion. Accordingly, a choice was made to not further discuss this model in Chapter 3, also to not further complicate our discussion on the key principles of PDAT that are necessary to understand the argumentation analysis performed. It may not be surprising that most attention went out to the *argumentative stage*, where arguments are put forward to support the standpoint. Some elements of the *opening stage* were also considered; in the ideal situation, major assumptions should be exchanged that explain why an expert interprets evidence the way he does, which PDAT refers to as *starting points*. Analyses of the *confrontation stage* and the *concluding stage* were not performed. In Chapter 4, the scope was limited to the *argumentation stage* altogether.

Given the *cross-sectional* rather than *longitudinal* nature of our analyses in Chapters 2 and 4 (see paragraph 5.3.4), it was not possible to assess the dynamic exchange of arguments between the studied parties over time. We are aware that the papers analyzed in Chapter 3 are embedded within an expansive, ongoing scientific discussion that may cover a wider range of topics than discussed in these respective publications. In Chapter 3, we argued that the piloting-nature of the study warranted the limited scope of the study. Another point is that the dynamic exchange in the EDS scientific debate is characterized by author groups that continuously vary and, more

importantly, that the specific scopes and aims of publications differs over time. The discussion in the studied papers revolves around the 2012 WHO-UNEP report, while other papers specifically address non-monotonic dose-response relationships or whether EDSs can be subjected to risk assessment. One would then expect that the same arguments re-occur in different papers of the same (or highly similar) author groups, while standpoints will differ from paper to paper, depending on the scope and aim, and a range of arguments specific to that scope will also emerge. It would be interesting to find out if future research could employ PDAT to corroborate our findings from Chapter 3 with the broader discussion on EDS science. However, such research should be attentive to the complexities that could arise when comparing the standpoints and argumentation contained in papers with different scopes and aims.

5.4.6 Experiences with using PDAT at RIVM

In general, PDAT requires the analysts to take up another role and view the exchange of argumentation under scrutiny from a much more abstract level than what an EDS expert would do. In this regard, PDAT has shown a tension between being a highly-trained expert and adopting a role of honest observer/analist. On the basis of their personal values, their scientific training and their ongoing activity in their field of expertise, any expert will develop value-based views about the evaluation of a risk issue, and what is necessary to bring the field forward. Even more so, all of the project members participate in complex scientific discussions on EDSs, or even have tasks to elucidate or defend, from an RIVM, Dutch or personal scientific position, certain views or standpoints in national or international gremia. As acknowledged by some of the members of this research project, precisely the resulting professional 'biases' make it difficult to take up a role as passive spectator or observer, so these deserve explicit attention when using PDAT.

Another tension lies with the areas of expertise of the involved project members. The theoretical and methodological underpinnings of PDAT are diametrically different, in almost every regard, from the underpinnings of the life sciences disciplines relevant to studying the risks posed by exposure to EDSs. This means that the RIVM project members brought in the toxicological, epidemiological or risk assessment expertise, and the members from the University of Amsterdam brought in their argumentation theory expertise. I was trained to have a sufficient proficiency in both fields, to allow me to perform the two empirical studies, but then again, I lacked the in-depth expertise in both areas to consider all intricacies. This required me to cooperate intensively with all involved project members, and similarly requires project members to cooperate with each other, though their areas of expertise, and with that all the concepts, terminology and scientific paradigms, are fundamentally different. An aspect of PDAT that potentially exacerbated confusion was that some of its concepts, such as 'standpoint', 'argument', 'argumentation', or 'difference of opinion', are named after terms that are also used in 'common language', but bear a highly specific and theory-laden meaning in the context of PDAT. This required extra consideration from the project members. Explicit consideration of this consequence of interdisciplinary work, and accordingly, a mutual initial investment to have at least a rudimentary understanding of each other's scientific disciplines is essential to adequately use PDAT.

Another point of attention is the relatively time-consuming and labor-intensive nature of the PDAT-based analyses. The process of identifying an argumentation structure of a text includes several steps: thoroughly understanding the content and context of the text, identifying the main arguments and standpoints and identify the logical links between the standpoints and arguments. Particularly the identification of logical links requires a meticulous, line-by-line analysis that is iterative by nature. It took chapter 3 co-author HvK about two months of work to complete the argumentation structures of the two analyzed publications. Chapter 4 co-author LM worked about 2,5 months to identify the argumentation structures of the 28 responses included in the analysis, while SCSC worked about another 2,5 months on the process of categorizing all identified arguments. Though PDAT allows for highly detailed and thorough scrutiny of the texts under consideration, a substantial time-investment should be incorporated in the project planning.

Overall, PDAT is invaluable as a method to 'read between the lines' in a maximally unbiased and impartial way. Chapter 3 shows that PDAT is suited to help identify implicit interpretative and normative elements in a highly technical and complex scientific discussion. Since the analytical framework of PDAT is topic-independent, PDAT could be used to study implicit interpretative or normative ambiguities that may occur in other scientific controversies. PDAT can also be used to gain insight in the range of key arguments that make up a discussion. In combination with a structured categorization procedure, this data could be used to identify the range of themes disputed among key stakeholders (see chapter 4). However, care should be taken to ensure that analyses are maximally unbiased and true to intended meaning of the text. The time-investment is substantial, so for the present PhD project with a relatively long lead time and room for trial-and-error, the investment was warranted given the achieved level of detail in the analyses. However, in the context of less flexible or more time-constrained projects, one should evaluate upfront whether the costs of thoroughly applying PDAT outweigh the benefits, and whether less meticulous analytical methods may be more efficient and, hence, more applicable.

5.5 IMPLICATIONS FOR FUTURE RESEARCH AND POLICY DEVELOPMENT

This study on scientific controversies in EDS has broken new ground by bringing together eight different frameworks from different scientific domains in an overarching framework and by applying argumentation analyses to explore the underlying argumentation of apparent controversies in EDS evaluation and regulation. Yet, it also only 'scratched the surface'.

In paragraph 5.3.5, suggestions were made on how de discussions on EDS science, policy and regulation could proceed. These suggestions followed from ideas and concepts provided by theoretical frameworks like the IRGC Risk Governance Framework and the NRC Analytic-Deliberative process. Likewise, diagnoses of the science system, such as 'Post-Normal Science', 'Mode-2 Science' and 'Technologies of Humility' also provide mostly theory-based ideal typical descriptions of the 'optimal' functioning of the scientific enterprise. Given the theory-ladenness of these ideas and frameworks, it might be worthwhile to assess how these hold up in the

(messy) every day practice of EDS subject-matter experts, expert policy advisors, policy makers, regulators and other key stakeholders. Such inclusive approaches will likely require thorough though flexible participation and decision-making processes, and ask for specific types of knowledge and experiences.

Accordingly, such future inclusive projects could explicitly focus on evaluations of the normative values held by these parties, or normative ambiguities occurring among these parties. Making explicit of the project's participants personal and professional values and assumptions could help unveil why specific differences of opinion have been difficult to settle. Procedures for science-based argumentation analyses, such as those provided by PDAT, could be used for this purpose, as shown in Chapters 3 and 4. First and foremost, adding to the project team (applied) ethicists or other professionals with knowledge and experience with values in science is recommended. It might also help to train project members in having an (at least) rudimentary understanding of how values in science are present, why it is important to acknowledge their existence and how one could adequately deal with them. Exposing them to literature introducing the topic of values in science, like 'A Tapestry of Values' of Elliott (2017) might also be an adequate starting point. On a more fundamental level, it may also be worthwhile to teach (university) students about values in science during courses that specifically focus on the role of their (developing) scientific expertise in modern society.

While the use of PDAT in this PhD research project may have been worthwhile, the accompanying analyses are time and labor intensive. The development of tools for argumentation analysis that may be faster, easier and simpler to use, would be encouraged. This would make the study of argumentation accessible to broad ranges of smaller projects. The effectiveness of other theories and models of argumentation, such as Toulmin's Model of Argumentation could also be piloted in future projects.

The present project only addressed a limited part of the overarching theoretical framework. It would be interesting to see whether it is indeed useful to apply it in its entirety in a research project, where contrasting advocacy coalitions are discerned, the range of (observable) arguments is identified, compared and contrasted, and the underlying (implicit) values are made explicit. Such a project could be performed with EDS as research topic, but it would also be interesting to study other high-profile environmental health issues, such as 5G electromagnetic fields, nitrogen deposition in (vulnerable) rural areas and COVID-19. In addition, the limits to the breadth of expertise of the RIVM members of this research project should be acknowledged, which was mostly life sciences oriented (particularly toxicology, epidemiology, risk assessment and environmental psychology). In future projects, it is encouraged to expand the project team using our theoretical framework with social scientists (e.g. experts in science technology studies, psychology and sociology of risk, policy sciences, public administration), and experts in humanities (e.g. scholars of argumentation theory, philosophy of science, (applied) ethics).

Future research could also study whether the findings of the two empirical studies could be corroborated in other studies when using the same approach (PDAT). In the context of Chapter 3, publications of other EDS author groups could be analyzed, to study whether the identified value differences can also be expanded to broader EDS scientific community, or whether these are specific to the two studied author groups. One could also assess whether different analytical approaches will yield similar results. For example, interviews (particularly Socratic dialogue style) with individual EDS experts could be performed to further explore different types of values held by experts, or to reflect on the values held by other experts. Focus groups (also dialogue style) could be used to have different experts reflect together on similarities and differences in the values held, and where these differences may originate from. With regard to Chapter 4, one could perform interviews or focus groups with the contact persons of the governmental entities and stakeholder organizations included in the study, to gain a more in-depth understanding of the interests, roles and positions on both sides of the spectrum of regulatory preferences.

In conclusion, this study has shown merits and some limitations of argumentation analysis of an ongoing scientific and regulatory controversy, exemplified with the EDS case. It can act as a catalyst to more inter- and transdisciplinary research in the area on other (environmental) health controversies. Hopefully, it will also lead to more awareness about the role of normative elements in scientific controversies in subject-matter experts, risk assessors, regulators, policy makers and representatives of NGO's and the public at large.

Appendices

A1.1 Example of interview protocol used during exploratory case study

INTRODUCTION (MAX. 5 MIN.)
[If required:]

- ➔ We specifically focus on differences in the governance of endocrine disrupting chemicals across EU member states.
- ➔ That is, the way EU member states deal with the risks posed by (potential) EDCs - such as BPA - differs significantly.
- ➔ Where some countries (e.g. France) seem to attain a more (pro)active regulatory approach in coping with EDC risks, other countries (e.g. the Netherlands) seem to attain a more passive regulatory approach, i.e. following the EC regulatory strategy.
- ➔ We are interested in the driving forces for the differences in these regulatory approaches.
- ➔ Specifically, we hypothesize that differences between EU member states' regulatory approaches may be (partially) explained by the role that specific stakeholders and their arguments have played in adopting some novel piece of a member state's EDS policy.
- ➔ (That is, there may be a wide range of arguments at stake, such as scientific arguments based on the risk assessment of BPA, social-ethical arguments regarding the fairness of risk/benefit distribution and voluntariness of exposure to BPA, politically-oriented arguments, etc..)
- ➔ (At this moment, we study the (French) proposal to restrict the use of BPA in thermal paper, though we aim to focus on other countries' policy interventions in the domain of EDC in the near future).
- ➔ (To get a balanced view of the background of the restriction proposal and the stakeholders involved, we wish to interview a wide variety of stakeholders from science, NGO's, industry and the (national) government.)

Thank you very much for your agreement to participate in this interview. Before I begin, first some procedural affairs I'd like to emphasize. I would like to **record** this interview, because I can then focus more strongly at listening. Recording also greatly improves the quality of the analysis. Do you agree with this?

I wish to emphasize that we will treat the resulting recording and ensuing materials as **confidential**, meaning that only directly involved project members will have access to this data. We will make a transcript of the recording which we will feedback to you to check for any factual mistakes that might have occurred. In the reporting of our results, we will make sure there are no quotes that could be traced to an individual interviewee. Therefore, not all

parts of this interview will be part of the formal analysis, some parts solely function to provide some background information.

Of course, there are no right or wrong answers in this, I am purely interested in your personal perspectives on the topics that we will discuss today. Do you have any questions about these procedural matters? If not, I would now like to start with asking the first question to start the interview.

BACKGROUND QUESTIONS (MAX. 5 MIN.)

- *Could you describe your position within RES?*
- *Could you briefly describe the mission of RES?*
- *Could you briefly describe the mission of RES with regard to endocrine disrupting substances?*

Now I would like to switch to the scientific aspects of BPA. Various assessments of the risks posed by exposure to BPA have been conducted. As you may know, ANSES published an extensive assessment of the risks associated to BPA exposure, as did EFSA.

SCIENCE OF BPA (MAX. 15 MIN.)

- *To what degree would you consider the scientific knowledge regarding BPA risks certain?*
  - To what degree would you consider this knowledge 'contested'?
- *Could you say something about the main risks associated to BPA given current exposure levels in the population?*

Now I would like to switch to the role of stakeholders, the general public's risk perception of BPA and the choice to restrict the use of BPA in specifically thermal paper. I would like to discuss the role of stakeholders specifically in the context of the restriction of BPA in thermal paper. I would like to discuss the general public's risk perceptions more in the context of general safety issues related to BPA.

RESTRICTION OF BPA IN THERMAL PAPER IN DETAIL (MAX. 30 MIN.)

- *Could you say something about your personal role, or the role of RES in the process of the restriction of BPA in thermal paper?*
- *Could you say something about the moment when specifically the use of BPA in thermal paper drew attention, or was problematized?*
  - What was the first moment that BPA in thermal paper was discussed?
  - What was the initiating moment of the process?
  - What were major issues of discussions?
  - Why were these events important?

- Was there one single event that has been most important?

- *Could you say something about the sequence of events that happened after that?*
  - *Societal discussion*
  - *Political reactions*
- *Could you say something about the stakeholders that played a major role in this sequence of events?*

- *Could you say something about the extent certain stakeholders played a major role in the decision to initiate the restriction process?*
  - What were their main positions?
  - What were their main arguments supporting these positions?

- *Could you briefly sketch an overall time-line of the major steps taken during the restriction process?*
  - What was the first moment that BPA in thermal paper was discussed?
  - What was the initiating moment of the process?
  - What were major issues of discussions?
  - Why were these events important?
  - Was there one single event that has been most important?

- *Could you say something about the current perceptions of the French general public regarding the safety of using thermal paper containing BPA (if any)?*

- *Could you say something about the current perceptions of the French general public regarding the safety of BPA in general (if any)?*
  - To what extent is the safety of BPA in general a topic of discussion in the French society?
  - To what extent is the safety of BPA in general a topic of discussion in the French society?

- *Do you have any insight into why specifically thermal paper was subjected to a restriction, in contrast to other applications of BPA?*
  - Possible reasons mentioned by interviewee (list not complete!):
    - Relatively redundant application with little societal benefits
    - Easily replaceable with less hazardous substitute
    - Protection of specific target groups (female cashiers of reproductive age)
    - Potential difficulties associated with limiting the use of other application of BPA (e.g. inner coating food-containing cans)

- *To what extent do you think the restriction was cost-effective in terms of reducing current risks to health and improving public health?*

- *To what extent do you feel was the restriction proposal supported by other countries?*

To finish up, I would like to discuss what other persons or institutions have been involved in the process leading to the restriction proposal.

FOLLOW-UP – GATHERING CANDIDATES FOR FUTURE INTERVIEWS (MAX. 5 MIN.)

- *Could you suggest a few people that may be relevant to talk to as well*?
    - Representatives from the policy domain?
    - NGO representatives?
    - Industry representatives?
    - Others?
- *Do you wish to add something with regard to the topics we've discussed?*

A3.1 Detailed description of pragma-dialectical argumentation theory

In order to systematically perform an argumentation analysis, we use approaches and insights provided by the pragma-dialectical argumentation theory (PDAT), developed by van Eemeren and colleagues (van Eemeren et al. 2010; van Eemeren and Grootendorst, 1984). According to PDAT, the goal of an argumentative discussion is solving a *difference of opinion* by means of acceptable *argumentation*. Indeed, the name of PDAT is inspired by its intellectual foundation; the study of pragmatics and the study of dialectics. Scholars in the field of pragmatics study how 'real life' communication and interactions attain meaning within a certain context (van Eemeren et al., 2014). For example, the phrase 'I can't get this computer working' could be interpreted in multiple ways. The speaker could simply assert that he cannot get the computer to work. Alternatively, the speaker could also implicitly request help in getting the computer to work. The exact intended meaning of the phrase may then depend on the specific situation, identity of the speaker, intonation of the utterance, body language, prior conversations and other contextual factors. Second, PDAT's dialectical orientation refers to solving differences of opinion in a *reasonable* manner (van Eemeren et al., 2014). In terms of the PDAT, *reasonableness* refers to advancing sound argumentative moves that do not impede resolving a *difference of opinion*. A typical example of *unreasonable argumentation* is putting forward *argumentation* that is based on an invalid reasoning scheme.

In this chapter, the argumentation analysis is performed in accordance with two vital steps of PDAT. The first step is the ideal model of a critical discussion (van Eemeren et al., 2010). This model consists of four stages that discussants should ideally follow when participating in a critical discussion. These stages are the *confrontation stage*, *opening stage*, *argumentation stage* and *concluding stage* (see Table A3.1). The key implication of the ideal model is that discussants should go through all stages to maximize chances of *reasonably* solving a *difference of opinion*. For example, when discussants do not go through the *confrontation stage*, they may not have agreement on the exact *standpoint* that should be under discussion. Then, a Babylonian confusion of tongues is lurking, making it practically impossible to resolve the intended

Table A3.1: The four stages of an ideal critical discussion

Confrontation stage	A *difference of opinion* is established, that is, one party's *standpoint* has been subjected to criticism or doubt by another party.
Opening stage	The parties engaged in the *difference of opinion* decide to resolve the conflict. They assign various roles (i.e., protagonist or antagonist) to one another and they establish whether there is enough common ground to build a discussion on
Argumentation stage	The protagonist(s) defend(s) his *standpoint* by putting forward *argumentation* and attempts to remove (anticipated) doubt and refute (possible) criticism from the opposing party.
Concluding stage	The extent of conflict resolution will be assessed, based on the acceptability of the *argumentation*. Ideally, the antagonist retracts the opposite *standpoint* (or doubt) if the protagonist has defended his *standpoint* successfully, or the protagonist retracts his *standpoint* if he fails to remove objections or doubt from the antagonist.

difference of opinion satisfactorily. In this paper, the *opening stage* is of particular interest. In the *opening stage*, the parties should explore what prerequisite knowledge is required to productively participate in the discussion, among others. In the ideal situation, the parties should come to an agreement about *starting points*. Note that a wide variety of *starting points* relevant to (a part of) the scientific debate of EDS science is discussed in the results section of the main body of this paper.

The second step pertains to argumentation stage of the ideal model of a critical discussion. Here, PDAT distinguishes between the external and internal organization of *argumentation*. The former is referred to as *argumentation structure,* whereas the latter is referred to as *argument scheme*. The *argumentation structure* provides a complete overview of the *standpoints* and all underlying *argumentation* and their hierarchal relationships. In the case of *complex argumentation* (i.e., when a *standpoint* is supported by more than one argument), three types of *argumentation* may be possible: *multiple argumentation*, *coordinative argumentation* and *subordinative argumentation* (van Eemeren et al, 2010), see also Table A3.2. Note that a*rgumentation structures* do typically not include every sentence of the analyzed exchange of argumentation, but only includes those propositions that the analyst interprets as *standpoints* or *argumentation*.

Table A3.2: The three types of complex argumentation.

Multiple argumentation	Multiple, (in principle) equally weighing arguments that are considered strong enough to individually support a *standpoint* successfully, forming a series of alternative defenses to the same *standpoint*. *Example: You couldn't have seen Peter drive a car last week, because Peter wasn't here last week and Peter doesn't have a driver's license.*
Coordinative argumentation	Multiple arguments that require each other to provide a conclusive defense, for example due to individually 'weak' arguments. *Example: You shouldn't buy that new computer, because it's too expensive and you don't have that kind of money.*
Subordinative argumentation	Multiple arguments that form a 'chain of arguments'; *argumentation* is put forward in support of other *argumentation*, instead of directly supporting a *standpoint*. *Example: I could not work yesterday, because I had severe stomachaches. I probably got these aches from eating spoiled fish.*

Argument schemes describe the specific relationship between a *standpoint* and the underlying argument. Generally, three types of *argument schemes* are discerned: *argumentation based on a symptomatic relation*, *argumentation based on a relation of analogy* and *argumentation based on a causal relation* (van Eemeren et al., 2010), see also Table A3.3. Of course, *argument schemes* frequently include premises that remain unexpressed. In terms of PDAT, *unexpressed premises* are considered not unintentionally omitted elements implicitly present in the *argumentation* for which a party can be held accountable (van Eemeren et al., 2010). The phrases between curved brackets in Table A3.3 are examples of identified *unexpressed premises*. For example, the *standpoint* 'I like your shirt' is linked with the underlying argument 'because it is purple'

through the *unexpressed premise* 'and being purple is considered characteristic for a likeable shirt'. Note that all three types of *argument schemes* include several sub variants, but these will not be discussed here (see e.g. van Eemeren et al., 2010). Finally, *critical questions* can be used to evaluate the appropriateness of the *argument scheme* applied in some *argumentation*. Note that an 'evaluation of appropriateness' requires an inherently normative judgment of the analyst.

Table A3.3: the three main types of argument schemes

Argumentation based on a symptomatic relation	An argument refers to a sign or typical feature that is relevant for the *standpoint*, thereby supporting it. *Example: I like your shirt, because it is purple* *(and being purple is considered characteristic for a likeable shirt)*
Argumentation based on a relation of analogy	A *(sub)standpoint* is defended by an argument that acknowledges a similarity or resemblance of something in the *standpoint*. *Example: Andy should get a discount, because Sam also got a discount* *(and Andy's situation is considered analogous to Sam's situation)*
Argumentation based on a causal relation	The argument put forward presents a cause or action of an effect mentioned in the *standpoint* to be defended *Example: You should wear a jacket outside, otherwise you might catch a cold (and not wearing a jacket may consequently cause one to catch a cold)*

A3.2.1 Example of a descriptive argumentation analysis

To illustrate the procedures pertaining to a descriptive argumentation analysis in practice, the *argumentation* structure of the text below will be reconstructed, one *unexpressed premise* will be made explicit and one *argument scheme* will be discussed, including the resulting *critical questions*.

"Cats make great pets. Here are some reasons why. First, cats don't need to take walks like dogs. Second, cats have soft, fluffy fur. […] Finally, if you have mice in your house […]. cats can catch them." (Teacherspayteachers, 2017)

The *argumentation structure* may be reconstructed as follows:

1. Cats make great pets
   1.1 They do not need to take walks like dogs
   1.2 Cats have soft, fluffy fur
   1.3 They can catch mice

Note that every *subordinative argument* is notated with an extra number one (.1) after the previous notation. *Multiple argumentation* is notated with whole numbers (1.1, 1.2, 1.3) whilst *coordinative argumentation* is notated with letters (1.1a, 1.1b, 1.1c). To illustrate making *unexpressed premises* explicit, the focus will be exclusively on the *standpoint* and the first argument: 'cats make great pets, because they don't need to take walks like dogs'.

Modus ponens

1. If p, then q (If cats don't need to take walks (like dogs), then they make great pets)
2. p (Cats don't need to take walks)
3. q (They make great pets)

This type of reasoning makes the *argumentation* logically valid, but it does not improve insight or add additional information. Instead of solely identifying this so-called *logical minimum*, a *pragmatic optimum* could be formulated , that is, the *pragmatically reasonable unexpressed premise* (van Eemeren and Grootendorst, 1992). An example would be 'taking your pet out for a walk is not a pleasurable activity'. Depending on the information that can be derived from the context, a list of *subordinative arguments* could be made explicit, such as: 'because it takes time to take them for a walk', 'it is harder to leave them on their own' or 'you don't always want to go out in bad weather'. If the context were to be different, the *subordinative arguments* would also change, which means that contextual information can provide clues that cannot be determined by the premise itself. Adding the suggested subordinative *unexpressed premises* would give the following *argumentation structure*:

1. Cats make great pets
  1.1 They don't need to take walks like dogs
    (*1.1') Taking your pet out for a walk is not a pleasurable activity*
    (1.1'1) It takes time to take them for a walk
    (1.1'2) It is harder to leave them on their own
    (1.1'3) You do not always want to go out in bad weather
1.2 Cats have soft, fluffy fur
1.3 They can catch mice

Based on the *unexpressed premise* (1.1'), the associated *argument scheme* may be categorized as *argumentation based on a symptomatic relation*, given the fact that cats do not need to be taken out on a walk is seen as characteristic or typical of the fact that they are great pets. For *argumentation based on a symptomatic relation*, the *critical questions* to evaluate the soundness of the *argument scheme* are derived (van Eemeren et al., 2010):

Y is true of X (being great pets is true of cats)
because: Z is true of X (not needing walks is true of cats)
and: Z is symptomatic of Y (not needing walks is symptomatic of being a great pet)

The *critical questions* to ask are:

- Aren't there also other non-Y's that have the characteristic Z?
- Aren't there also other Y's that do not have the characteristic Z?

In the previous example that would be:

- Aren't there also other pets that are not great that do not need walks?
- Aren't there also other great pets that do need walks?

Similar to the formulation of the *pragmatic optimum*, the extent to which a satisfactory answer is found for these *critical questions* depends on the available contextual information. In this example, the protagonist may have an earlier stated opinion on other types of (domestic) animals and the degree of 'active' maintenance the preferred animals should require.

A4.1 Identification of option preferences

Organisation	Specific wording of policy preference	Derived option preference
BCPC	… **additional elements of [hazard] characterization** which relate to severity and reversibility of effect, resulting from rates of absorption, distribution, metabolism, and excretion, should also be included in this assessment [next to potency]	Including potency, and additional elements of hazard characterization
CEFIC	… a single set of criteria for the determination of endocrine disrupting properties, which uses the WHO/IPCS definition as a basis, but which also takes into account the **relevance of the adverse effect** (that is: severity of effect, (ir)reversibility of effect, potency and lead toxicity) ensures the correct identification of the relevant substances	Including potency, and additional elements of hazard characterization
ECPA	… ECPA believes that **all elements of hazard characterization** should be included into the criteria, that is: severity of effect, (ir)reversibility of effect, potency and lead toxicity should all be incorporated into the final ED criteria (i.e. potency should not be the only consideration to characterize the hazard)*	Including potency, and additional elements of hazard characterization
PlasticsEurope	… **additional elements of hazard characterization** should be an integral part of the optimal policy options. As such, severity of effect, (ir)reversibility of effect, potency and lead toxicity should be considered as part of the weight of evidence approach	Including potency, and additional elements of hazard characterization
BEUC	**We support option 3** as a good way forward – *the first two categories should be used for regulation. The third category (potential) is important to request industry to gather more information on the potentially harmful properties*	Category approach
HEAL	HEAL believes **Option 3 best captures & enables** the optimal use of the existing state of the art science in a way that will best serve protection of public health	Category approach
PAN Europe	**Creating classes [i.e. Option 3] is the best option** … *- this Option includes elements of the whole WHO definition but also two 'dangerous' elements – human relevance and that the effects should occur in the absence of other toxic effect*	Category approach
Endocrine Society	The Endocrine Society supports the scientific approach outlined in Options 2 and 3. However, **we prefer Option 3** due to certain weaknesses in Option 2 …	Category approach

EurEau	… we **strongly support option 3** of the proposed roadmap …	Category approach
Belgium	We consider that an Endocrine disruptor has to be identified based on the IPCS/WHO 2002 definition … - **it would be useful** to have a supplementary list for 'suspected ED' or 'endocrine active substances' … - *this should be a dynamic list, not used for classification purposes as such*	Different variation of category approach
France	The French authorities are **in favour of the option 3**	Category approach
Netherlands	From the regulatory perspective, we **consider it useful** to define a single category of 'potential endocrine disrupters', which can be used to trigger further dedicated testing on a case by case basis	Different variation of category approach
UK	… while the UK's fundamental objection to the use of hazard criteria to ban the use of chemicals under some legislation would remain, Option 4 **appears to be the least inappropriate – all aspects of hazard identification and hazard characterization** should be taken into account in identifying EDs for regulatory purposes**	Including potency, and additional elements of hazard characterization
Australia	Consistent with our trade obligations and international best practice, Australia uses a **risk-based rather than a hazard-based** approach … - … potency should not be the only additional consideration; other criteria including specificity, severity and irreversibility **should be considered** …***	Risk-based option
Canada	None of the four scenario's proposed in the EU roadmap for impact assessment of ED criteria includes a risk-assessment option. Therefore, Canada's view is that none of the options presented in this roadmap **are consistent with the [risk-based] approach** … ***	Risk-based option
New Zealand	… New Zealand **does not support any of the four proposed options** – Because New Zealand uses a risk-based rather than a hazard-based approach for assessing the safety of chemicals, none of the options presented are directly comparable – Option 4 is an improvement over the other options presented	Risk-based option
US	No specific preference mentioned, multiple appeals to risk-based approaches*** - Creating technical regulations on the basis of hazard based criteria are often (1) more trade restrictive than necessary because risk based mitigation measures exist …	Indirectly stated preference for risk-based option

Austria (AGES/UBA)	No explicit preference stated for one of the four options	No specific preference
Denmark (EPA/VFA)	**We support** option 3	Category approach
Finland (TUKES)	The Finnish CA **supports** the option 3	Category approach
Germany (BfR)	From a scientific point of view this option [Option 4] **is supported** for human health risk assessment****	Including potency
Germany (UBA)	For **environmental aspects only Option 3 is suitable** - No specific Option supported related to human health	Different variations of category approach
Hungary (NICS)	No explicit preference stated for one of the four options – note: Options 1, 2 and 4 are (somewhat implicitly) considered unfavorable, though no explicit support for Option 3 is observed	No specific preference
Sweden (KEMI)	The Swedish Chemicals Agency (KEMI) **supports** option 3	Category approach
Health Canada	Health Canada uses a **risk-based rather than a hazard-based** approach for assessing the safety of chemicals***	Risk-based option
Norway FSA	**We support** option 3 …	Category approach

*ECPA ultimately prefers a full risk assessment approach, but since the current regulatory decision making does not such approach, their priority is on ensuring sound regulatory criteria, before returning to a complete risk assessment framework

**UK raises the possibility to allow risk assessment of at least some EDCs, in accordance with defined criteria

*** The EU is not explicitly required or insisted to follow a risk assessment approach, but these countries' own risk assessment approach to substance regulation (including EDC regulation) is consistently repeated, including the requirements to adhere to 'risk-based' approaches established by the WTO

****The BfR response refers to a study conducted by the BfR, in which also additional elements of hazard characterization have been included. Nevertheless, in the BfR response, there is specific, explicit support for Option 4 'as is' (i.e., no explicit mention of support for an 'extended Option 4' type of option)

*****The UBA response refers to the official German position paper issued earlier, in which an 'extended Option 4' type of option is advocated. However, there is no explicit support for any Option related to human health, though the exercise performed by BfR (in which Option 4 was found most favorable) is referred to

A4.2 Overview of identified themes and topics, including all associated arguments

***Theme* 1: Mechanistic understanding of EDSs**

Issue 1.1: Timing of exposure and effects

◄ The timing of exposure, particularly during vulnerable periods of development, is relevant for EDCs – France (1), Swedish KEMI (4), Norwegian FSA (5), HEAL (7), PAN Europe (8), Endocrine Society (9)

◄ The activity of EDCs depends on age – PAN Europe (8)

◄ The exposure to EDCs during early development can cause adverse effects later in life – HEAL (7), Endocrine Society (9)

◄ The exposure to EDCs during pregnancy can cause adverse effects in children/offspring – HEAL (7), Endocrine Society (9)

Issue 1.2: Dose-response

◄ Hormones and EDCs are active at very low concentrations – France (1), Endocrine Society (9),

◄ Adverse effects can occur at very low doses of exposure to EDCs – Swedish KEMI (4), Norwegian FSA (5), HEAL (7), Endocrine Society (9)

◄ EDCs can exert non-monotonic dose-response behavior - Swedish KEMI (4), Norwegian FSA (5)

◄ EDCs can exert non-threshold behavior – Swedish KEMI (4), PAN Europe (8)

◄ Humans and the environment may be exposed to high levels of weakly potent EDCs, which may pose a risk – Danish EPA (2), Norwegian FSA (5)

◄ A single hormone will have different effects at different times and places in the body during development and with different sensitivity – Endocrine Society (9)

◄ One chemical could affect multiple endocrine pathways - Endocrine Society (9)

◄ EDCs vary in their activity in different parts of the body and in different hormone systems – Swedish KEMI (4), HEAL (7), PAN Europe (8)

◄ Endocrine systems are often linked – Endocrine Society (9)

◄ The relationship between hormone concentration and gene expression can be influenced by very small changes in co-factor concentration and/or hormone receptor concentrations – Endocrine Society (9)

◄ The activity of hormones or EDCs on different endpoints can vary strongly – Endocrine Society (9)

► All substances, including oxygen and water, can cause adverse effects including death, if administered at sufficiently high doses; it is the dose that makes the poison – BCPC (16)

► Chemicals are expected to have the same effect at the same concentration in the same situation – United States (23)

► Different substances can have widely differing endocrine activities – New Zealand (22)

► Substances with very low potency may not be expected to have adverse effects – New Zealand (22)

► It is important to recognize the concentration or dose at which endocrine effects occur – UK (15)
► There is an extensive literature about 'endocrine disruption' as a general phenomenon, which includes claims about non-monotonic dose response and low dose effect, which do not appear to be relevant to the well-characterized mechanisms which have been examined for phthalates and brominated flame retardants (BFRs) – Australia (20)
► The actual hazard of a substance depends on its actual uses/applications and the subsequent nature/level of exposure as occurring in the 'real world' – United States (23)
► Results obtained in the laboratory may not be the same as those occurring when the substance is used in a non-laboratory environment, given the complexity of the endocrine system – Australia (20)

Issue 1.3: Mixture effects
◄ People are exposed to a mixture of EDCs – Norwegian FSA (5), HEAL (7), Endocrine Society (9)
◄ EDCs in a mixture could act together – Norwegian FSA (5)
◄ The exposure to a mixture of EDCs could lead to additive effects –Swedish KEMI (4), HEAL (7)
◄ Multiple EDCs could affect a single endocrine mode of action – Endocrine Society (9)
◄ Mixtures of EDCs should be studied, rather than single substances – Endocrine Society (9)

Issue 1.4: Assessment of environmental EDSs
◄ Interspecies variations are particularly relevant for endocrine modes of action – German UBA (13),
◄ Sensitivities for EDCs can differ across species - Danish EPA (2), Norwegian FSA (5), HEAL (7), PAN Europe (8), Hungarian NICS (26)
► Environmental hazards (including those applying to the aquatic compartment) cannot be assessed in a meaningful way without taking into account properties that affect an organism's exposure to a substance – UK (15)
► Whether mammalian effects are relevant in an environmental context depends on the properties that affect an organism's exposure – UK (15)

***Theme* 2: Regulatory considerations related to the identification of EDSs**

Issue 2.1: Availability of EDS-related data
◄ Sensitive EDC endpoints are missing in most regulatory testing guidelines, particularly effects following from exposure during sensitive periods, general early life exposures and late on-set effects – Danish EPA (2), Swedish KEMI (4), HEAL (7), The Netherlands (12), New Zealand (22)
◄ There is variability in the inclusion of EDC sensitive endpoints among studies – Danish EPA (2)

◄ There is a general lack of knowledge regarding endocrine modes of action/endocrine system of humans and other mammals, particularly during developmental stages – PAN Europe (8)
◄ Many endocrine modes of action/parts of the endocrine system (particularly those that do not involve estrogen, androgen, thyroid or steroid activity) are not elicited in most regulatory testing guidelines – Swedish KEMI (4), HEAL (7), Belgium (11)
◄ Mechanistic data is not included in the regulatory assessment of EDCs – PAN Europe (8)
◄ Current regulatory testing guidelines were designed for the assessment of adversity, not to reveal modes of action – The Netherlands (12)
◄ A single study involving exposure through the complete life cycle of a mammal, from conception to old age is missing – The Netherlands (12)
◄ A single study involving developmental exposure with follow-up into old age is missing – The Netherlands (12)
◄ There is a lack of knowledge on exposure during sensitive periods of development – Belgium (11)
◄ There are data gaps concerning background levels – Belgium (11)
◄ There are data gaps concerning the identification of key events in the different adverse outcome pathways – Belgium (11)
► The required toxicology database is generally very extensive for pesticide active constituents – ECPA (18), Australia (20)
► For the human health situation, current testing and assessment methodologies are generally suitable to derive dose/concentration levels which can be considered safe, provided substances have undergone comprehensive evaluation – German UBA (13), German BfR (14),
► Tests have already been designed to identify some ED mode of action as well as some ED adverse effects, of which the OECD conceptual framework on ED provides useful information on these aspects, though limitations remain – Belgium (11)

Issue 2.2: Quality and variability of data on EDSs
◄ There are strong variations in the available EDSs test data and its quality – The Netherlands (12)
◄ Old data were obtained using tests that covered only a small number of ED-parameters, whereas newer, more sensitive tests have been updated to cover more and more sensitive ED-related parameters – Swedish KEMI (4), The Netherlands (12)
◄ Tests can have different sensitivities in terms of statistical power – Swedish KEMI (4),
◄ There is a tendency to use less animals in shorter duration studies, and to include not all relevant and sensitive parameters, which reduces the power of the studies – The Netherlands (12)
◄ There is variability of data sets and study designs in terms of exposure at sensitive windows of exposure – Danish EPA (2)
◄ There is variability of data sets and study designs in terms of different sensitivities due to group sizes – Danish EPA (2)
◄ Regulatory assessments should also include human epidemiological data and other (non-in vivo) model systems – Endocrine Society (9)

► There is no convincing evidence to assume that levels of uncertainty are generally different regarding EDCs as compared to other toxic substances, although absolute certainty regarding safe dose/concentration levels is generally not achievable – German UBA (13), German BfR (14)

Issue 2.3: Establishing causality

◄ Given the limitations of currently required regulatory testing, it is nearly impossible to (unequivocally) prove a causal relationship between hormonal changes and adverse health effects (in an experimental study) – BEUC (6), The Netherlands (12)

◄ Ubiquitous exposure to EDCs combined with the multi-factorial nature of related conditions makes it difficult to unequivocally establish causation – Danish EPA (2), Endocrine Society (9)

◄ Data that demonstrates a causal link between a chemical exposure and an effect are generally only available after many years or decades of research – Australia (20)

◄ Adverse effects must be observed in intact organisms, thus the WHO/IPCS definition does not allow identification based on in vitro testing – Norwegian FSA (5)

► Signals of endocrine activity do not automatically imply adverse effects in intact animals – Australia (20)

► Changes in hormone levels accompanying adverse health effects are not proof of causality per se, since hormonal imbalance may be secondary to the adverse effects observed – The Netherlands (12)

► In addition to endocrine disruption, there are other biochemical and physiological mechanisms which can be affected by excessive chemical exposure – Australia (20)

► There are many mechanisms apart from the direct receptor binding sometimes assumed for EDCs – Australia (20)

► Endocrine interactions are often just one of a multitude of factors contributing to an adverse outcome – Australia (20)

► Natural stressors can cause modulation of endocrine systems, with potential adverse effects – CEFIC (17)

► Physiological stresses such as parasitism, temperature, hypoxia, caloric intake and food restriction are known to induce adverse endocrine effects – CEFIC (17)

► Physical interactions that can alter endocrine functions include environmental conditions such as temperature, salinity, and/or pH – CEFIC (17)

► In many toxicological studies, hormones may well change to due stress, due to adaptation, and due to chemical action, such changes can also be potentially assumed to be a plausible link to observed effects – CEFIC (17), PlasticsEurope (19)

► Substances should only be considered as EDCs of regulatory concern when there are clear adverse effects unambiguously caused by a well identified and empirically described mode of action – CEFIC (17), PlasticsEurope (19)

► The OECD Conceptual Framework should be used to support robust weight-of-evidence evaluations, combined with formal frameworks for assessing adversity, modes of action, human relevance and causation (such as the WHO/IPCS mode of action/human relevance framework , evolved from the WHO/IPCS 2002 report, which were both adapted from the Hill considerations – CEFIC (17), PlasticsEurope (19)

► Regulatory assessments should differentiate between scientifically robust and weak studies – CEFIC (17), PlasticsEurope (19)
► Assessments of the risks of substances for top predators, primates and other larger animals should be based on substantial data that warrants a concern – CEFIC (17), PlasticsEurope (19)

Issue 2.4: Assessment of environmental EDSs
◄ There is a particular lack of knowledge on endocrine modes of action/endocrine system of non-human/non-mammal vertebrate and invertebrate species/environmental EDCs – PAN Europe (8), Belgium (11)
◄ For major taxa, no adequate testing exists to derive safe dose/concentration levels – German UBA (13)
◄ Standard testing methods only monitor very severe adverse effects – German UBA (13)
◄ Lack of tests to cover whole life cycle for all relevant taxa – Swedish KEMI (4), Belgium (11)

***Theme* 3: Consistency with existing regulatory frameworks**

Issue 3.1: Consistency with PPP/BP Regulations
◄ The PPPR/BPR wording of 'may cause' should be acknowledged in the criteria/level of evidence required to fulfil the criteria should correspond to 'may cause' wording in PPPR and BPR – Swedish KEMI (4), BEUC (6), HEAL (7), PAN Europe (8)
◄ The PPPR and BPR wording of 'may cause' requires inclusion of 'potential EDC' WHO/IPCS definition/requires lower degree of causality than WHO/IPCS wording of 'consequently causes' – Danish EPA (2), BEUC (6), HEAL (7), PAN Europe (8)
◄ The PPPR and BPR intend to ban both 'confirmed' and 'potential' EDCs –HEAL (7), PAN Europe (8)
◄ Implementation of the ultimate criteria in the PPPR and BPR is facilitated by adopting Option 3, since the provisions between the PPPR and BPR differ – HEAL (7)
► The PPPR requires exposure considerations, and thus, a broader risk assessment approach, since it stipulates that substances with ED-properties (that may cause adverse effects in humans) cannot be approved for use unless human exposure (under realistic conditions of use) is negligible – Australia (20)
► BPR requires a broader risk assessment approach since it stipulates that substances considered having ED-properties (that may cause adverse effects in humans) should not be approved unless the risk to humans is negligible – Australia (20)
► Regulatory resources under the PPPR should be directed at chemicals whose hazardous properties might conceivably hold a threat to humans or the environment under real-world conditions – UK (15)
► Under PPPR, the European Commission is required to develop a single set and specific criteria, not a series of categories – BCPC (16), ECPA (18)
► PPPR and BPR do not have provisions for multiple categories, only for ['confirmed'] EDCs and do not foresee multiple categories – CEFIC (17), PlasticsEurope (19)

► Categorization is not required to request nor to generate additional data on possible endocrine-related effect to reach regulatory decisions [within the PPPR], relevant legislative provisions are already in place and are not limited to endocrine related effects but are more expansive – ECPA (18)

Issue 3.2: Consistency with REACH Regulation

◄ 'Probable serious effects' wording of REACH relates to lower degree of causality than the wording of the WHO/IPCS definition for a 'confirmed' EDC ('consequently causes') – Danish EPA (2)

► REACH process of substance evaluation (i.e. CoRAP) should suffice for substances for which the status is unclear, a multiple category system is not required – CEFIC (17), PlasticsEurope (19)

► The need to differentiate [using potency/fate] is recognized in REACH, where substances with ED-properties can be added to the candidate list for authorization. This process goes beyond the WHO/IPCS definition alone, given the wording 'substances … for which there is scientific evidence of probable serious effects to human health or the environment which give rise to an equivalent level of concern to those of other [CMR/PBT/vPvB] substances and which are identified on a case-by-case basis' – UK (15)

Issue 3.3: Consistency with CLP Regulation

◄ Since many known ED effects are carcinogenic or toxic to reproduction/are of equivalent concern as CMR substances, and carcinogenic and reprotoxic effects are not classified on the basis of potency, EDCs should also not be classified on the basis of potency - Danish EPA (2), HEAL (7)

◄ Since characterization of substances as carcinogenic, mutagenic and/or reprotoxic occurs on the basis of weight of evidence, EDCs should also be regulated on the basis of weight of evidence considerations – Danish EPA (2)

◄ Categorizing EDCs in accordance with the weight of evidence is consistent with current approaches to rank other chemicals in the EU (CLP) and under GHS – BEUC (6), HEAL (7)

◄ Including potency for EDCs could cause some substances to be identified as carcinogenic or reprotoxic, but not as EDC – Danish EPA (2)

► CMR adverse effects are well-defined toxicological outcomes which are suitable to categorization, while ED is not an adverse effect in itself – CEFIC (17), ECPA (18), PlasticsEurope (19)

► Assessment of aquatic toxicity uses chronic NOEC values to distinguish between different degrees of long-term hazards, and if CLP regulation sees value in distinguishing between different levels of potency towards the aquatic environment, then such discrimination is surely relevant for the threat of ED in the aquatic environment. Similar arguments can be made for human health hazard considerations – UK (15)

► When endocrine modes of action are the mechanistic background of adverse effects that will lead to CLP classification (such as reprotoxic, carcinogenicity, target organ toxicity, environmental hazard), doublecounting in hazard identification may occur, which may lead to

unnecessary administrative burden in regulatory decision making, will have quenstionable value in terms of risk management and thus requires urgent attention – Finnish TUKES (3)

Issue 3.4: Consistency with Cosmetics Regulation

◄ Cosmetics do not necessarily need to be permanently listed under Category 3 of Option 3, since the wording of Category 2 ('experimental studies') could also include in vitro studies and in vitro studies could be used to place a cosmetic substance in either Category 3 or 2, provided that an advanced battery of in vitro studies is available – Norwegian FSA (5)

► Category III would become a 'category of no way out' for cosmetic substances, due to the ban on animal testing in the field of cosmetics. It might be impossible to clarify whether these substances are really EDCs – Austria (AGES/UBA) (25)

Issue 3.5: Usefulness of categorizing EDSs as a specific regulatory substance category

◄ Endocrine disruption is not a specific endpoint but a network of mechanisms that lead to differential endocrine-related diseases, where strong and weak triggers on specific sites may equally result in the development of disease – PAN Europe (8)

► Endocrine disruption is not an adverse end-point per se, but rather a generic terminology artificially grouping a collection of different modes of action and different adverse effects of variable nature, severity and concern – CEFIC (17), ECPA (18), PlasticsEurope (19), Australia (20)

► Use of a single umbrella term obscures that there are many different mechanisms apart from the direct receptor binding sometimes assumed for EDCs; further elucidation of this issue come from the OECD's Adverse Outcome Pathway framework, which treats all components of the endocrine system on their own merits – Australia (20)

► Regulation should be focused on adverse effects of impact on the endocrine system (as function of the dose administered and further related toxicokinetics), not on the mode of action (which may, or may not, result in adverse effects without reference to the risk of such adverse effects actually occurring in practice) – BCPC (16), ECPA (18)

► Observed adversity is related most directly to damage to organ systems (e.g. in case of thyroid hormones, the affected organ is the developing brain); disruption of the endocrine system may be transient, while damage to the organ system is ongoing – Australia (20)

► Australian chemical regulatory agencies do not consider that ED is an adverse end-point per se, but rather is a mode of action potentially leading to other (eco)toxicological outcomes, such as reproductive, developmental, carcinogenic or ecological effects, which are routinely considered in reaching regulatory decisions for pesticide active constituents – Australia (20)

► Genetic modification, the process of genetic modification of crops, is not per se a source of hazard or risk, yet 'GM crops' are – wrongly – perceived by many to be dangerous – BCPC (16)

***Theme* 4: Evaluation of specific issues related to a category approach**

Issue 4.1: Practicality of applying a category approach

▲Option 3 enables the needed differentiated regulatory classification according to the level of scientific evidence available – France (1), Danish EPA (2), Finnish TUKES (3), BEUC (6), HEAL (7), Endocrine Society (9),

▲ Option 3 acknowledges the level of evidence and scientific uncertainties related to EDCs – Norwegian FSA (5), Australia (20)

▲ Option 3 is the best approach to make use of the information from the currently available test methods – Swedish KEMI (4)

▲ Option 3 enables identification of 'potential' EDCs on the basis of screening, in vitro testing or QSAR methods – Norwegian FSA (5)

▲ Using only the 'confirmed' EDC WHO/IPCS definition does not systematically or effectively consider the available scientific evidence - HEAL (7)

▲ Option 3 enables inclusion of further advances in science and scientific knowledge – Norwegian FSA (5), Endocrine Society (9)

▲ Option 3 enables identification of knowledge gaps for substances that could be EDCs – PAN Europe (8)

▲ Option 3 enables the development and systematic application of integrated testing strategies – Danish EPA (2)

▲ Option 3 enables (transparent) comparative ranking of substances, in accordance with the level of scientific evidence available – BEUC (6), HEAL (7)

▲ Option 3 enables flagging of those substances for which a conclusive determination has not been made/which require additional information – Swedish KEMI (4), Austria (AGES/UBA) (25)

▲ Option 3 enables a system to request further information, by flagging of substances or allocating substances to a (dynamic) list, to request further information and subsequently classify these substances in accordance with one of the three categories, or remove the concern – Denmark EPA (2), Norwegian FSA (5), Belgium (11), German UBA (13)

▲ Option 3 enables differentiated regulatory action/prioritized regulatory action for certain substances, in terms of (further) dedicated testing, assessment and management of substances and in accordance with the available weight of evidence – Danish EPA (2), Finnish TUKES (3), Swedish KEMI (4), BEUC (6), HEAL (7), Netherlands (12)

▲ Option 3 enables differentiated allocation of regulatory resources for further investigation and, if necessary, testing of the most problematic substances – Danish EPA (2)

▲ Option 3 enables a system of 'early warnings' triggers for industry and pesticide/biocide producers to provide more information to demonstrate the safety of substances, consider substances with due prudence or disregard or replace potentially hazardous substances –, Finnish TUKES (3), Swedish KEMI (4), PAN Europe (8), German UBA (13)

▲ Option 3 enables better analysis when examining the costs & benefits of different possible regulatory actions as part of impact assessments, and when considering new policy measures – HEAL (7)

▲ Having categories facilitates decision making, based on experience from the assessment of CMR substances – Norwegian FSA (5)
▼ It is difficult to identify an EDC from a potential EDC – Australia (20)
▼ The evidence to categorize chemicals that are in some manner 'suspected' to be EDCs is 'insufficiently strong' – UK (15)
▼ Categorization of EDCs has no (scientific) basis, since ED artificially groups together substances with variable modes of action and adverse effects under the same terminology – CEFIC (17), ECPA (18), PlasticsEurope (19), Australia (20)
▼ No specific criteria to differentiate between the 'levels of evidence' of the three categories have been proposed – Endocrine Society (9)
▼ Insufficient clarity about the distinction between Category I and II, as currently proposed – Australia (20), New Zealand (22)
▼ Based on the proposed [imprecise] criteria, it is impossible to determine if substances are Endocrine Disruptors – CEFIC (17), PlasticsEurope (19)
▼ Unclear which legal consequences follow from the assignment to the different categories – German BfR (14), UK (15), PlasticsEurope (19), Austria (AGES/UBA) (25)
▼ Unclear what the purpose and fitting of the categories is in the context of the PPPR and BPR – CEFIC (17), PlasticsEurope (19)
▼ Unclear who decides if and when a substance moves from one category to another – UK (15)
▼ When categories are included, the current legal texts in which EDCs are mentioned need to be changed – Austria (AGES/UBA) (25)
▼ Introduction of additional categories based on weight of evidence considerations might lead to difficulties in including different legal consequences of the different assignments to the categories into the respective legal texts – Austria (AGES/UBA) (25)
▼ Adopting [imprecise] categories would create a major unpredictability and lack of consistency for industry, in relation to more fragmented chemical regimes – CEFIC (17), PlasticsEurope (19)
▼ Categories may lead to unnecessary stigmatization, blacklisting or tainting by association of listed substances – UK (15) CEFIC (17), ECPA (18), PlasticsEurope (19), Austria (AGES/UBA) (25)
▼ 'Black lists' will be highly vulnerable to misinterpretation, misuse and unwarranted additional primary and secondary regulation – ECPA (18)
▼ Category III is very broad – Australia (20)
▼ Category III is too broad – New Zealand (22)
▼ Unclear why Category III, as holding position, is needed – UK (15)
▼ Terminology of 'Endocrine Active Substances' in Category III is inaccurate; in common understanding, EAS do not usually include harmful consequences, in contrast to the description provided in the Roadmap – Austria (AGES/UBA) (25), Hungarian NICS (26)
▼ Considered unlikely that it will be possible to perform further assessment of many of the substances included under Category III – New Zealand (22)
▼ Category approach can be potentially very confusing in terms of stakeholder communication – UK (15)

▼ Dose applied and exposure levels of relevant organisms and their physiological systems are not included – BCPC (16)

Issue 4.2: Anticipated impact on the categorization of substances

▲A wider range of substances with EDC properties will be captured [than with the other options] – PAN Europe (8)

▼ Substances will be labeled 'suspected' EDCs, despite the fact that they have been fully evaluated under the relevant regulatory schemes – CEFIC (17), ECPA (18), PlasticsEurope (19)

▼ A very large number of substances could be classified purely on the basis of an in vitro finding that may not be of relevance for human health or the environment –Australia (20), New Zealand (22)

▼ The categorization under Option 3 is not helping to identify substances of regulatory concern – CEFIC (17), PlasticsEurope (19)

▼ Substances categorized under Category III may be subject to some form of restriction or concern for a very long time without adequate scientific justification, and in many cases unnecessarily – New Zealand (22)

Issue 4.3: Anticipated impact on expert judgment processes

▲Reduces pressure on assessors to have to make forced dichotomous 'yes/no' decisions – HEAL (7), Australia (20)

▲Preempts differentiation in expert judgments based on different preferences regarding 'precautionary identifications' of EDCs – HEAL (7)

▲Enables assessors to appropriately incorporate uncertainty (i.e. 'to judge the varying strength of evidence') – Australia (20)

▼ Weight-of-evidence assessments, which require detailed knowledge and experience, may generate different outcomes if performed by different regulators – German BfR (14)

▼ Based on current definitions of the categories, they require extensive interpretation by experts, leading to different categorization of the same substance – CEFIC (17), PlasticsEurope (19), Australia (20), New Zealand (22)

Issue 4.4: Anticipated impact on the amount of animals used in animal testing

▲Option 3 meets the requirement in the PPPR to reduce the number of animal studies, by including in vitro studies – Norwegian FSA (5)

▲Defining a single category of 'potential' EDCs will reduce animal use and costs – The Netherlands (12)

▼ Additional animal testing would be required to be able to clarify the status of active substances placed in Categories 2 and 3, which would involve the sacrifice of large numbers of animals to provide unnecessary data which would not add any additional understanding to the toxicological behavior of pesticide active substances – ECPA (18)

▼ Risk assessment would limit the need for additional, unnecessary animal testing – ECPA (18)

Issue 4.5: Suitability of a category approach for dealing with environmental EDS

▲Option 3 is the only suitable option for the environmental situation – German UBA (13)

***Theme* 5: Evaluation of specific issues related to including potency**

Issue 5.1: Practicality of including potency

▼ Potency cannot account for different endocrine mechanisms of action/different activity at different sites of the endocrine system – Swedish KEMI (4), HEAL (7), PAN Europe (8)
▼ Potency thresholds are impossible to establish for critical windows of susceptibility during early development – Swedish KEMI (4), PAN Europe (8)
▼ Potency cannot account for non-threshold mechanisms – Swedish KEMI (4)
▼ Potency of a hormone or EDC on different endpoints differ [significantly] – Endocrine Society (9)
▼ Potency does not take into account non-linear dose-response curves – Swedish KEMI (4), Norwegian FSA (5)
▼ Potency does not take into account low-dose effects – Swedish KEMI (4), Norwegian FSA (5)
▼ Potency cannot account for additive effects of mixtures – Swedish KEMI (4)
▼ Inclusion of potency will potentially introduce underestimation of mixture toxicity - Danish EPA (2)
▼ Potency differs across age – PAN Europe (8)
▼ Potency cannot take into account the effects on populations and vulnerable subgroups – Swedish KEMI (4)
▼ Potency does not take into account the variability of included sensitive endpoints between tests – Denmark EPA (2)
▼ Potency does not take into account the variability of inclusion of exposures at sensitive periods between tests – Denmark EPA (2)
▼ Potency does not take into account different sensitivities due to different group sizes used in tests – Denmark EPA (2)
▼ Schemes based on cut-off values for potency are scientifically indefensible/unjustifiable/arbitrary and too formulaic to accommodate the subtleties needed for scientifically sound judgments – Danish EPA (2), Swedish KEMI (4), BEUC (6)
▼ For the human health situation, establishing potency values would primarily be a policy decision, these could not be scientifically determined – Belgium (11)
▼ Unclear which cut-off values would be used – Austria (AGES/UBA) (25)
▼ Scientifically speaking, potency is not part of hazard identification, or the identification of substances as EDCs – HEAL (7)
▼ Potency is not part of hazard identification but could play a role at a later stage: in hazard characterization or in an eventual risk assessment process – Danish EPA (2), Finnish TUKES (3), The Netherlands (12)
▼ Including potency is not truly hazard-based and designating an EDC should be purely hazard-based – Swedish KEMI (4), BEUC (6), HEAL (7), The Netherlands (12),

▼ Inclusion of potency would lead to lower protection of human health and the environment – HEAL (7)
▼ Old substances, tested a long time ago, will be favored over new substances, simply because the new substances were tested using more sensitive testing methods, thereby showing higher potency – The Netherlands (12)
▲Including potency has the benefit of being properly scientifically-based, unlike the other options presented – BCPC (16)
▲ It makes no sense to regulate stringently substances that are of insufficient potency (unless they have the potential to transform to more hazardous substances in the environment) - UK (15)
▲ If one substance is readily biodegradable and of very low potency, and another is persistent and of high potency, it is clear which is of most concern – UK (15)
▲ Potency and fate in the environment are essential to inform on level of concern, whether or not there is specific exposure information – UK (15)
▲ Substances with very low potency may not require classification or restriction, since such substances may not be expected to have adverse effects – New Zealand (22)
▲ Including potency enables ranking of substances – New Zealand (22)
▲ Including potency enables differentiation of substances of various potencies – German BfR (14)
▲ Including potency helps prioritize substances to be regarded as of higher or lower regulatory concern –German UBA (13), German BfR (14), New Zealand (22)
▲ Option 4 enables differentiation of regulatory action, including further research – German BfR (14), New Zealand (22)
▲ For the human health situation, Option 4 scored high on sensitivity for hazard identification – German UBA (13), German BfR (14)
▲ For the human health situation, reproducibility of Option 4 was comparable to the option based on hazard identification only – German UBA (13), German BfR (14)
▲ Facilitating decision making is regarded as better than the hazard identification option, due to integration of elements of hazard characterization – German UBA (13), German BfR (14)
▲ Option 4 recognizes the importance of concentration, dose and exposure levels of organisms and their physiological systems in determining whether a substance may cause adverse effects – BCPC (16)
▲ Potency is a key determinant of whether a substance may induce adverse effects at toxicologically and/or environmentally relevant concentrations – BCPC (16), CEFIC (17), ECPA (18), PlasticsEurope (19)
▲ Potency is critical for comparisons to exposure for risk-based methods – Australia (20)

Issue 5.2: Anticipated consequences of including potency
▼ Potential to wrongly omit important chemicals, on the basis of 'low potency', for identification as EDC and subsequent regulatory action – BEUC (6), HEAL (7), PAN Europe (8)
▼ Low potent substances with high exposures could not be identified, and could lead to a risk for humans or the environment – HEAL (8)

▼ Weak-potent EDCs acting together will not be identified - Danish EPA (2), HEAL (8)

Issue 5.3: Suitability of including potency for dealing with environmental EDS

▼ Potency differs across species, since species have different sensitivities for specific EDCs – Danish EPA (2), Norwegian FSA (5), HEAL (7), PAN Europe (8), Hungarian NICS (26)

▼ For the environmental situation, appropriate potency cut-off values could not be scientifically determined - Belgium (11)

▼ Difficult to select taxa that provide the basis for potency cut-off value establishment – HEAL (7)

▼ Including potency is unworkable for the environment – Danish EPA (2)

▼ Only highly potent substances used on the taxa in the tests considered will be identified – HEAL (7)

A4.3 BRIEF DISCUSSION OF OPTIONS 1 AND 2

Since maintaining the 'interim criteria' (Option 1) and adopting the WHO/IPCS 2002 definition of an EDS 'as is' (Option 2) have been evaluated predominantly negatively, we will shortly describe the most recurring substantive objections to these options. A detailed treatise of the arguments for and against these option can be found in EC (2015). With regard to arguments challenging the acceptability of maintaining the 'interim criteria', these could be subdivided into three groups, that each consist of one or more arguments:

One argument relates to the limited scientific basis of maintaining the 'interim criteria':

- Various respondents consider this option to be inconsistent with the widely supported WHO/IPCS definition for the identification of EDSs, since Option 1 does not explicitly consider endocrine modes of action or requires that the mode of action considered subsequently causes adverse effects

Two arguments relate to the identification of substances when the 'interim criteria' are maintained:

- Some respondents refer to the potential for type I errors, since non-carcinogenic and non-reprotoxic substances will unlikely be identified, potentially leading to false negatives
- Other respondents refer to the potential for type II errors, since substances could be misidentified as EDSs, which do categorize as CMRs, but do not work through an endocrine mode of action, potentially leading to false positives

Four arguments relate to the regulatory impracticality of maintaining the 'interim criteria':

- The lack of a formal definition for the term 'toxic to endocrine organs' is noted, leading to potential regulatory inconsistencies and uncertainty
- It is argued that the interim criteria only apply to human health, not to the environmental, while it is intended that the ultimate criteria are protective of both human health and the environment
- The regulatory value of doublecounting substances as both CMRs and EDSs is considered to be questionable by some respondents
- It is argued that the interim criteria complicate the (future) harmonization of provisions for EDSs in all relevant regulatory frameworks; the interim criteria are only formally articulated in the PPPR and BPR, but in slightly different wordings, and the current provisions for EDSs in other relevant regulations (REACH, Cosmetics regulation) also differ from one another

With regard to the arguments challenging the acceptability of adopting the WHO/IPCS 2002 definition of an EDS 'as is', these mostly relate to shortcomings that are remedied by either a category approach (Option 3) or the inclusion of potency, in addition to the adoption of the WHO/IPCS definition (Option 4). The category approach and the inclusion of potency both build on Option 2, but contain different additional provisions that, for various reasons, appeal to different respondents. In general, the WHO/IPCS definition is widely supported as an

appropriate definition to define EDSs for scientific purposes. However, this definition is widely considered as too limited to be used directly in a regulatory setting. We do not discuss these limitations in more detail here, since these can be inferred from our detailed descriptions of the respondents' evaluations of the category approach and the inclusion of potency.

References

ADAS, 2011. *Evaluation of the benefits provided by the azole class of compounds in wheat, and the effect of losing all azoles on wheat and potato production in Denmark, France and the UK. Report 1 – Impact of the loss of all azoles*. Wolverhampton, UK: ADAS.

Autrup, H, Barile, FA, Blaauboer, BJ, Degen, GH, Dekant, W, Dietrich, D, Domingo, JL et al. 2015. Principles of Pharmacology and Toxicology Also Govern Effects of Chemicals on the Endocrine System. *Toxicological Sciences, 146* (1):11–15. doi:10.1093/toxsci/kfv082.

Aven, T, & Renn, O. 2010. *Risk management and governance: Concepts, guidelines and applications*. Berlin Heidelberg: Springer-Verlag.

Bakir, V. 2005. Greenpeace v. Shell: Media exploitation and the social amplification of risk framework (SARF). *Journal of Risk Research*, *8*(7–8), 679–691. https://doi.org/10.1080/13669870500166898

Beausoleil, C. et al., 2016. Review of non-monotonic dose-responses of substances for human risk assessment. *EFSA supporting publication 2016: EN-1027*.

Bergman, Å, Andersson, AM, Becher, G, van den Berg, M, Blumberg, B, Bjerregaard, P, Bornehad, CG, et al. 2013. Science and policy on endocrine disruptors must not be mixed: a reply to a "common sense" intervention by toxicology journal editors. *Environment Health 12(69)*.

Bergman, Å, Becher, G, Blumberg, B, Bjerregaard, P, Bornman, R, Brandt, I, Casey, SC et al. 2015. Manufacturing Doubt about Endocrine Disrupter science - A Rebuttal of Industry-Sponsored Critical Comments on the UNEP/WHO Report "State of the Science of Endocrine Disrupting Chemicals 2012. *Regulatory Toxicology and Pharmacology, 73* (3):1007–1017. doi:10.1016/j.yrtph.2015.07.026.

Beronius, A, Rudén, C, Hakansson, H, & Hanberg, A. 2010. Risk to all or none? A comparative analysis of controversies in the health risk assessment of Bisphenol A. *Reproductive Toxicology*, *29*, 132–146. https://doi.org/10.1016/j.reprotox.2009.11.007

Borgert, CJ, Matthews, JC, Baker, SP. 2018. Human-relevant potency threshold (HRPT) for EAα agonism. *Archives of Toxicology 92*: 1685-1702.

Carson, R. 1962. *Silent Spring*. Boston, MA: Houghton Mifflin.

CDC (Centers for Disease Control and Prevention). 2020. *About DES*. URL: https://www.cdc.gov/des/consumers/about/index.html

Clahsen, SCS, van Kamp, I, Hakkert, BC, Vermeire, TG, Piersma, AH and Lebret, E. 2019. Why do countries regulate environmental health risks differently? A theoretical perspective. *Risk Analysis 39 (2)*: 439-461.

Clahsen, SCS, van Klaveren, HS, Vermeire, TG, van Kamp, I, Garssen, B, Piersma, AH and Lebret, E. 2020. Understanding conflicting views of endocrine disruptor experts: a pilot study using argumentation analysis. *Journal of Risk Research 23 (1)*: 62-80.

Colborn, T, Clement, C. 1992. *Chemically-Induced Alterations in Sexual and Functional Development: The Wildlife/Human Connection*. Princeton, NJ: Princeton Scientific Publishing Co., Inc.

Colborn, T, Dumanoski, D and Myers, JP. 1996. *Our Stolen Future: Are We Threatening Our Fertility, Intelligence, and Survival? A Scientific Detective Story*. Dutton, New York.

Covello, VT, & Merkhofer, MW. 1993. *Risk assessment methods: Approaches for assessing health and environmental risks*. New York: Plenum Press.

Dake, K. 1992. Myths of nature: Culture and the social construction of risk. *Journal of Social Issues*, *48*(4), 21–37. https://doi.org/10.1111/j.1540-4560.1992.tb01943.x

Dietrich, DR, von Aulock, S, Marquardt, H, Blaauboer, B, Dekant, W, Kehrer, J, Hengstler, J et al. 2013. Scientifically unfounded precaution drives European Commission's recommendations on EDC regulation, while defying common sense, well-established science and risk assessment principles. *Chemico-Biological Interactions 205*(1): A1-A5.

Dietrich, DR. 2016. Don't mar legislation with pseudoscience. *Nature 535*: 355.

Dolan, T, Howsam, P, Parsons, DJ, & Whelan, MJ. 2013. Is the EU drinking water directive standard for pesticides in drinking water consistent with the precautionary principle? *Environmental Science and Technology*, *27*(10), 4999–5006. https://doi.org/10.1021/es3049555g

Douglas, H. 2000. Inductive Risk and Values in Science. *Philosophy of Science 67* (4):559–579.

Douglas, M. 1966. *Purity and danger: An analysis of the concepts of pollution and taboo*. London and New York: Routledge.

Douglas, M. 1970. *Natural symbols: Explorations in cosmology*. London: Cresset Press.

Douglas, M. 1992. *Risk and Blame - Essays in Cultural Theory*. London: Routledge.

Douglas, M., & Wildakvsy, A. 1983. *Risk and culture: An essay on the selection of technological and environmental dangers*. Berkeley and Los Angeles, CA: University of CA Press.

EFSA (European Food Safety Authority). 2015. *Conclusion on the peer review of the pesticide risk assessment of the active substance glyphosate*. Parma, Italy: European Food Safety Authority. URL: https://www.efsa.europa.eu/en/efsajournal/pub/4302.

Elliott, KC. 2017. *A Tapestry of Values: An Introduction to Values in Science*. New York: Oxford University Press.

EP (European Parliament). 2019. *Endocrine Disruptors: from Scientific Evidence to Human Health Protection*. URL: https://www.europarl.europa.eu/RegData/etudes/STUD/2019/608866/IPOL_STU(2019)608866_EN.pdf

EPA (Environmental Protection Agency). 2016. *How EPA regulates drinking water contaminants*. URL: https://www.epa.gov/dwregdev/how-epa-regulates-drinking-watercontaminants.

EPA (Environmental Protection Agency). 2020. *The NRC risk assessment paradigm*. URL: https://www.epa.gov/fera/nrc-risk-assessment-paradigm.

EC (European Commission). 2014. *Roadmap – Defining criteria for identifying Endocrine Disruptors in the context of the implementation of the Plant Protection Product Regulation and Biocidal Products Regulation*. URL: https://ec.europa.eu/health/endocrine_disruptors/process_en

EC (European Commission). 2015a. *Report on – Public consultation on defining criteria for identifying endocrine disruptors in the context of the implementation of the Plant Protection Product Regulation and Biocidal Products Regulation*. URL: https://ec.europa.eu/health/sites/health/files/endocrine_disruptors/docs/2015_public_consultation_report_en.pdf

EC (European Commission). 2015b. DG GROW Briefing Note. URL: https://www.asktheeu.org/en/request/5226/response/17030/attach/7/3%20050417%20Briefing%20Plastics%20Europe%200n%20BPA%202%20Redacted.pdf

EFSA (European Food Safety Authority), 2013. Scientific Opinion on the hazard assessment of endocrine disruptors: Scientific criteria for identification of endocrine disruptors and appropriateness of existing test methods for assessing effects mediated by these substances on human health and the environment. *EFSA Journal 2013; 11*(3):3132.

FAO/WHO (Food and Agriculture Organization of the United Nations/World Health Organization). 2016. *Pesticide residues in food 2016—Special session of the joint FAO/WHO meeting on pesticide residues*. Rome, Italy: Food and Agriculture Organization of the United Nations/World Health Organization. URL: https://www.fao.org/3/a-i5693e.pdf.

FAO (Food and Agriculture Organization of the United Nations). 2020. *Hazard identification & characterization*. URL: http://www.fao.org/pesticide-registration-toolkit/registration-tools/assessment-methods/method-detail/en/c/1186994/.

Fischhoff, B, Slovic, P, Lichtenstein, S, Read, S, & Combs, B. 1978. How safe is safe enough? A psychometric study of attitudes towards risks and benefits. *Policy Sciences, 9*(2), 127–152. https://doi.org/10.1007/BF00143739

Frewer, LJ, Miles, S, & Marsh, R. 2002. The media and genetically modified foods: Evidence in support of social amplification of risk. *Risk Analysis*, *22*(4), 701–711. https://doi.org/10.1111/0272-4332.00062

Funtowicz, SO, & Ravetz, JR. 1993. Science for the post-normal age. *Futures* 25(7), 739–755. https://doi.org/10.1016/0016-3287(93)90022-L

Gibbons, M. 1999. Science's new social contract with society. *Nature, 402*: C81-C84.

Gibbons, M, Limoges, C, Nowotny, H, Schwartzman, S, Scott, P, Trow, M. 1994. *The new production of knowledge*. London, UK: SAGE Publications Ltd.

Graven, C, Dang, Z, Piersma, AH, Hakkert, B. 2016. *Endocrine disrupting chemicals in the EU legal frameworks – Human health perspective.* Bilthoven, the Netherlands: RIVM. URL: https://www.rivm.nl/bibliotheek/rapporten/2016-0137.pdf

Grendstad, G, & Selle, P. 2000. Cultural myths of human and physical nature: Integrated or separated? *Risk Analysis*, *20*(1),27–39.

Guillette, LJ Jr., Gross, TS, Masson, GR, Matter, JM, Percival, HF, Woodward, AR. 1994. Developmental abnormalities of the gonad and abnormal sex hormone concentrations in juvenile alligators from contaminated and control lakes in Florida. *Environmental Health Perspectives 102*(8), 680-688.

Herbst, L, Ulfelder, H, Poskanzer, DC. 1971. Adenocarcinoma of the Vagina – Association of Maternal Stilbestrol Therapy with Tumor Appearance in Young Women. *New England Journal of Medicine 284*, 878-881.

Hermans, MA., Fox, T, & van Asselt, MBA. 2012. Risk governance. In S. Roeser, R.Hillerbrand, P. Sandin, & M. Peterson (Eds.), *Handbook of risk theory*. The Netherlands: Springer.

Hisschemöller, M, and Hoppe, R. 1995. Coping with Intractable Controversies: The Case for Problem Structuring in Policy Design and Analysis. *Knowledge and Policy: The International Journal of Knowledge Transfer and Utilization, 8* (4):40–60.

Hofstede, G. 1980. *Culture's consequences: International differences in work-related values*. Beverly Hills: Sage Publications.

Hofstede, G. 1991. *Cultures and organizations: Software of the mind*. London: McGraw-Hill.

Hofstede, G. 2011. Dimensionalizing cultures: The Hofstede model in context. *Online Readings in Psychology and Culture, 2*(1). https://doi.org/10.9707/2307-0919.1014

Hofstede, G and Hofstede, GJ. 2005. *Cultures and organizations: Software of the mind*. New York: McGraw-Hill.

Hofstede, G, Hofstede, GJ, & Minkov, M. 2010. *Cultures and organizations: Software of the mind: Intercultural cooperation and its importance for survival*. New York: McGraw-Hill.

Hood, C, Rothstein, H, & Baldwin, R. 2001. *The government of risk: Understanding risk regulation regimes*. Oxford: Oxford University Press.

Hoppe, R. 2007. Applied cultural theory: Tool for policy analysis. In F. Fischer, G.J. Miller, & M.S. Sidney (Eds.), *Handbook of public policy analysis: Theory, politics, and methods*. Boca Raton, FL: CRC Press.

Hoppe, R. 2009. Scientific advice and public policy: Expert advisers' and policymakers' discourses on boundary work. *Poiesis & Praxis, 6*(3–4), 235–263. https://doi.org/10.1007/s10202-008-0053-3

Horel, S., CEO (Corporate Europe Observatory). 2015. *A Toxic Affair – How the chemical lobby blocked action on hormone disrupting chemicals*. Paris/Brussels: Stéphane Horel and Corporate Europe Observatory.

IARC (International Agency for Research on Cancer). 2017. *Some organophosphate insecticides and herbicides*. Lyon, France: International Agency for Research on Cancer. URL: https://monographs.iarc.fr/ENG/Monographs/vol112/mono112.pdf.

IPCS (International Programme on Chemical Safety). 2004. *IPCS risk assessment terminology*. Geneva: World Health Organization. URL: https://www.who.int/ipcs/methods/harmonization/areas/ipcsterminologyparts1and2.pdf?ua=1

IRGC (International RiskGovernance Council). 2005. *Risk governance: Towards an integrative approach*. Geneva: International Risk Governance Council.

IRGC (International Risk Governance Council). 2006. *White paper on nanotechnology risk governance*. Geneva: International Risk Governance Council.

IRGC (International Risk Governance Council). 2008. *An introduction to the IRGC risk governance Framework*. Geneva: International Risk Governance Council.

IRGC (International Risk Governance Council). 2009. *Risk governance of synthetic biology—Concept note*. Geneva: International Risk Governance Council.

IRGC (International Risk Governance Council). 2017. *Introduction to the IRGC risk governance framework, revised version*. Lausanne: EPFL International Risk Governance Center.

Jann, W, & Wegrich, K. 2007. Theories of the policy cycle. In F. Fischer, G.J. Miller,& M.S. Sidney (Eds.), *Handbook of public policy analysis: Theory, politics, and methods*. Boca Raton, FL: CRC Press.

Jasanoff, S. 1990. *The fifth branch*. Cambridge, MA: Harvard University Press.

Jasanoff, S. 2003. Technologies of Humility: Citizen Participation in Governing Science. *Minerva, 41*, 223-244.

Jenkins-Smith, HC, Nohrstedt, D, Weible, CM, & Sabatier, PA. 2014. The advocacy coalition framework: Foundations, evolution and ongoing research. In P.A. Sabatier & C.M. Weible (Eds.), *Theories of the policy process*. Boulder, CO: Westview Press.

Jenkins-Smith, HC, Silva, CL, Gupta, K, & Ripberger, JT. 2014. Belief system continuity and change in policy advocacy coalitions: Using cultural theory to specify belief systems, coalitions, and sources of change. *Policy Studies Journal*, *42*(4), 484–508. https://doi.org/10.1111/psj.12071

Johannesson, M, Hansson, SO, Rudén, C, & Wingborg, M. 1999. Risk management—The Swedish way(s). *Journal of Environmental Management*, *57*(4), 267–281. https://doi.org/10.1006/jema.1999.0300

JRC (Joint Research Centre), 2013. *Key scientific issues relevant to the identification and characterization of endocrine disrupting substances. Report of the Endocrine Disrupters Expert Advisory Group (ED EAG)*. Luxembourg: Publications Office of the European Union.

Kasperson, JX, Kasperson, RE, Pidgeon, N, & Slovic, P. 2003. The social amplification of risk: Assessing fifteen years of research and theory. In N. Pidgeon, R.E. Kasperson, & P. Slovic (Eds.), *The social amplification of risk*. Cambridge: Cambridge University Press.

Kasperson, RE, Renn, O, Slovic, P, Brown, HS, Emel, J, Goble, R, . . . Ratick, S. 1988. The social amplification of risk: A conceptual framework. *Risk Analysis*, *8*(2), 177–187. https://doi.org/10.1111/j.1539-6924.1988.tb01168.x

Klinke, A, Renn, O. 2002. A New Approach to Risk Evaluation and Management: Risk-Based, Precaution-Based, and Discourse-Based Strategies. *Risk Analysis, 22*(6), 1071-1094.

Kortenkamp, A, Martin, O, Faust, M, Evans, R, McKinlay, R, Orton, F, Rosivatz, E. 2011. *State of the art assessment of endocrine disrupters – Final Report*. URL: https://ec.europa.eu/environment/chemicals/endocrine/pdf/sota_edc_final_report.pdf
Kortenkamp, A, Martin, O, Evans, R, Orton, F, McKinlay, R, Rosivatz, E, Faust, M. 2012. Response to A critique of the European Commission Document, "State of the Art Assessment of Endocrine Disruptors" by Rhomberg and colleagues – letter to the editor. *Critical Reviews in Toxicology 42*(9): 787-789.

Lamb, JC, Boffetta, P, Foster, WG, Goodman, JE, Hentz, KL, Rhomberg, LR, Staveley, J, et al. 2014. Critical Comments on the WHO-UNEP State of the Science of Endocrine Disrupting Chemicals - 2012. *Regulatory Toxicology and Pharmacology, 69* (1):22–40. doi:10.1016/j.yrtph.2014.02.002.

Lebret, E. 2016. Integrated environmental health impact assessment for risk governance purposes; across what do we integrate? *International Journal of Environmental Research and Public Health*, *13*(71), 1–15. https://doi.org/10.3390/ijerph13010071

Löfstedt, RE. 2011. Risk versus hazard—How to regulate in the 21st century. *European Journal of Risk Regulation*, *2*(2), 149–168. https://doi.org/10.1017/S1867299X00001033

Löfstedt, RE, & Renn, O. 1997. The Brent Spar controversy: An example of risk communication gone wrong. *Risk Analysis*, *17*(2), 131–136. https://doi.org/10.1111/j.1539-6924.1997.tb00852.x

Matthiessen, P. 2003. Historical perspective on endocrine disruption in wildlife. *Pure and Applied Chemistry 75*(11-12). 2197-2206.

Markey, CM, Rubin, BS, Soto, AM, Sonnenschein, C. 2003. Endocrine disruptors: from Wingspread to environmental developmental biology. *Journal of Steriod Biochemistry & Molecular Biology 83*, 235-244.

Nohynek, GJ, Borgert, CJ, Dietrich, D, Rozman, KK. 2013. Endocrine disruption: Fact or urban legend? *Toxicology Letters, 223*, 295-305.

Norden. 2014. *The Cost of Inaction – A Socioeconomic analysis of costs linked to effects of endocrine disrupting substances on male reproductive health*. Copenhagen, Denmark: Nordic Council of Ministers.

Nowotny, H, Scott, P Gibbons, M. 2001. *Re-Thinking Science: Knowledge and the Public in an Age of Uncertainty*. Cambridge, UK: Polity Press.

NRC (National Research Council). 1983. *Risk assessment in the federal government: Managing the process*. Washington, DC: National Academy Press.

NRC (National Research Council). 1996. *Understanding risk: Informing decisions in a democratic society*. Washington, DC: National Academy Press.

NRC (National Research Council). 2009. *Science and decisions: Advancing risk assessment*. Washington, DC: National Academy Press.

Pielke, RAJ. 2007. *The honest broker: Making sense of science in policy and politics*. Cambridge: Cambridge University Press.

Piersma, AH, Hernandez, LG, van Benthem, J, Muller, AJJ, van Leeuwen, RFX, Vermeire, TG, van Raaij, MTM. 2011. Reproductive toxicants have a threshold of adversity. *Critical Reviews in Toxicology, 41*(6): 545-554.

PlasticsEurope, 2019. *Socio-economic Contribution.* URL: http://bisphenol-a-europe.org/socio-economic-contribution/.

Ravetz, JR. 1999. What Is Post-Normal Science. *Futures, 31*: 647–653.

Rayner, S. 1992. Cultural theory and risk analysis. In S. Krimsky & D. Golding (Eds.), *Social theories of risk*. Westport, CT: Praeger.

Redmayne, M. 2016. International policy and advisory response regarding children's exposure to radio frequency electromagnetic fields (RF-EMF). *Electromagnetic Biology and Medicine*, 35(2). https://doi.org/10.3109/15368378.2015.1038832

Renn, O. 2008. *Risk governance: Coping with uncertainty in a complex world*. London: Earthscan.

Rhomberg, LR, Goodman, JE, Foster, WG, Borgert, CJ, Van Der Kraak, G. 2012. A critique of the European Commission Document, "State of the Art Assessment of Endocrine Disruptors". *Critical Reviews in Toxicology 42*(6): 465-473.

Rijk, I., van Duursen, M., van den Berg, M. 2016. *Health costs that may be associated with Endocrine Disrupting Chemicals*. Utrecht, The Netherlands: IRAS.

RIVM (Dutch National Institute for Public Health and the Environment). 2016. Endocrine Disrupting Chemicals in the EU Legal Frameworks - Human Health Perspective. Bilthoven, The Netherlands: National Institute for Public Health and the Environment. URL: http://www.rivm.nl/Documenten_en_publicaties/Wetenschappelijk/Rapporten/2016/oktober/Endocrine_disrupting_chemicals_within_EU_legal_frameworks_human_health_perspective.

Rothstein, HF, Demeritt, D, Paul, R, Beaussier, AL, Wesseling, M, Howard, M, . . . Bouder, F. 2017. Varieties of risk regulation in Europe: Coordination, complementarity and occupational safety in capitalist welfare states. *Socio-Economic Review*, 1–28. https://doi.org/10.1093/ser/mwx029

Sabatier, PA. 1998. The advocacy coalition framework: Revisions and relevance for Europe. *Journal of European Public Policy*, 5(1), 98–130. https://doi.org/10.1080/135017688800000051

Sabatier, PA, & Jenkins-Smith, H. 1993. *Policy change and learning: An advocacy coalition framework*. Boulder, CO: Westview.

Sage, C, Carpenter, D, and Hardell, L. 2016. Comments on SCENIHR: Opinion on Potential Health Effects of Exposure to Electromagnetic Fields. *Bioelectromagnetics, 37* (3):190–484.

Sarewitz, D. 2000. *Science and environmental policy: an excess of objectivity*. In: Frodeman, R & Baker, VR (eds.), Earth Matters: The Earth Sciences, Philosophy, and the Claims of Community. Prentice-Hall. pp. 79-98.

Sarewitz, D. 2004. How science makes environmental controversies worse. *Environmental Science and Policy*, *7*, 385–403. https://doi.org/10.1016/j.envsci.2004.06.001

SCENIHR (Scientific Committee on Emerging Newly Identified Health Risks). 2015. Opinion on Potential Health Effects of Exposure to Electromagnetic Fields. *Bioelectromagnetics, 36* (6):480–484.

SCHER/SCENHIR/SCCS (Scientific Committee on Health and Environmental Risks/Scientific Committee on Emerging and Newly Identified Health Risks/Scientific Committee on Consumer Safety). (2013). *Making risk assessment more relevant for risk management*. Brussels: European Commission.

Schlumpf, M, Cotton, B, Conscience, M, Haller, V, Steinmann, B, Lichtensteiger, W. 2001. In vitro and in vivo estrogenicity of UV screens. *Environmental Health Perspectives, 109*(3), 239-244.

Science. 2016. *Why Europe may ban the most popular weed killer in the world*. URL: http://www.sciencemag.org/news/2016/06/why-europe-may-ban-most-popular-weed-killer-world.

Schwartz, M, & Thompson, M. 1990. *Divided we stand: Redefining politics, technology and social choice*. Philadelphia: University of Pennsylvania Press.

Siegrist, M, Keller, C, & Kiers, HAL. 2005. A new look at the psychometric paradigm of perception of hazards. *Risk Analysis*, *25*(1), 211–222. https://doi.org/10.1111/j.0272-4332.2005.00580.x

Simon, HA. 1957. *Models of man: Mathematical essays on rational human behavior in a social setting*. New York: John Wiley & Sons.

Slovic, P. 1987. Perception of risk. *Science*, *236*, 280–285. https://doi.org/10.1126/science.3563507

Slovic, P. 1999. Trust, emotion, sex, politics, and science: surveying the risk-assessment battlefield. *Risk Analysis*, *19*(4), 689–701. https://doi.org/10.1111/j.1539-6924.1999.tb00439.x

Slovic, P, Fischhoff, B, & Lichtenstein, S. 1976. Cognitive processes and societal risk taking. In J.S. Carroll & J.W. Payne (Eds.), *Cognition and social behavior*. London: Lawrence Erlbaum Associates, Inc.

Slovic, P, Fischhoff, B, & Lichtenstein, S. 1982. Why study risk perception? *Risk Analysis*, *2*(2), 83–93. https://doi.org/10.1111/j.1539-6924.1982.tb01369.x

Spruijt, P. 2016. *Expert views on scientific policy advice on complex environmental health issues*. PhD dissertation, Utrecht University, Utrecht.

Spruijt, P., Knol, AB, Petersen, AC, & Lebret, E. 2015. Different roles of electromagnetic field experts when giving policy advice: An expert consultation. *Environmental Health*, *14*(7), 1–10. https://doi.org/10.1186/1476-069X-14-7

Spruijt, P, Knol, AB, Petersen, AC, & Lebret, E. 2016. Differences in views of experts about their role in particulate matter policy advice: Empirical evidence from an international expert consultation. *Environmental Science and Policy*, *59*, 44–52. https://doi.org/10.1016/j.envsci.2016.02.003

Spruijt, P, Knol, AB, Torenvlied, R, & Lebret, E. 2013. Different roles and viewpoints of scientific experts in advising on environmental health risks. *Risk Analysis*, *33*(10), 1844–1857. https://doi.org/10.1111/risa.12020

Spruijt, P, Knol, AB, Vasileiadou, E, Devilee, D, Lebret, E, & Petersen, AC. 2014. Roles of scientists as policy advisers on complex issues: A literature review. *Environmental Science & Policy*, *40*, 16–25. https://doi.org/10.1016/j.envsci.2014.03.002

Steg, L, and Sievers, I. 2000. Cultural Theory and Individual Perceptions of Environmental Risks. *Environment and Behavior* 32 (2):250–269. doi: 10.1177/00139160021972513.

Tansey, J, & O'Riordan, T. 1999. Cultural theory and risk: A review. *Health, Risk and Society*, *1*(1), 71–90. https://doi.org/10.1080/13698579908407008

Teacherspayteachers. 2017. *Opinion writing mentor texts in second grade: what is the best pet?*. URL: https://www.teacherspayteachers.com/Product/Opinion-Writing-Mentor-Texts-in-Second-Grade-What-is-the-Best-Pet-600888.

Toulmin, SE. 1958. The Uses of Argument. Cambridge University Press.

Tversky, A, & Kahneman, D. 1973. Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, *5*, 207–232. https://doi.org/10.1016/0010-0285(73)90033-9

Tversky, A, & Kahneman, D. 1974. Judgment under uncertainty: Heuristics and biases. *Science*, *185*(4157), 1124–1131. https://doi.org/10.1126/science.185.4157.1124

van Asselt, MBA, & Renn, O. 2011. Risk governance. *Journal of Risk Research*, *12*(4), 431–449. https://doi.org/10.1080/13669877.2011.553730

van Eemeren, FH, Garssen, B, Krabbe, E, Henkemans, AFS, Verheij, B, and Wagemans, JHM. 2014. *Handbook of Argumentation Theory*. Dordrecht: Springer.

van Eemeren, FH, and Grootendorst, R. 1984. *Speech Acts in Argumentative Discussions: A Theoretical Model for the Analysis of Discussions Directed Towards Solving Conflicts of Opinion*. Dordrecht: Foris Publications.

van Eemeren, FH, Grootendorst, R. 1987. Fallacies in pragma-dialectical perspective. *Argumentation 1*: 283-301.

van Eemeren, FH, Grootendorst, R. 1992. *Argumentation, Communication and Fallacies: a pragma-dialectical perspective.* Hillsdale, NJ: Lawrence Erlbaum Associates.

van Eemeren, FH, Grootendorst, R, and Henkemans, AFS. 2010. *Argumentation: Analysis, Evaluation, Presentation*. New York: Routledge.

Vandenberg, LN, Colborn, T, Hayes, TB, Heindel, JJ, Jacobs, DR Jr., Lee, DH, Shioda, T, et al. 2012. Hormones and Endocrine-Disrupting Chemicals: Low-Dose Effects and Nonmonotonic Dose Responses. *Endocrine reviews 33*(3), 378-455.

Varret, C, Beronius, A, Bodin, L, Bokkers, BGH, Boon, PE, Burger, M, De Wit-Bos, L, et al. 2018. Evaluating the evidence for non-monotonic dose-response relationships: A systematic literature review and (re-)analysis of *in vivo* toxicity data in the area of food safety. *Toxicology and Applied Pharmacology, 339*: 10-23.

Weible, CM, & Sabatier, PA. 2007. A guide to the advocacy coalition framework. In F. Fischer, G.J. Miller, & M.S. Sidney (Eds.), *Handbook of public policy analysis: Theory, politics, and methods*. Boca Raton, FL: CRC Press.

Weible, CM, Sabatier, PA, & McQueen, K. 2009. Themes and variations: Taking stock of the advocacy coalition framework. *Policy Studies Journal*, *37*(1), 121–140. https://doi.org/10.1111/j.1541-0072.2008.00299.x

Weiss, C. 2003. Scientific uncertainty and science-based precaution. *International Environmental Agreements: Politics, Law and Economics*, *3*(2), 137–166. https://doi.org/10.1023/A:1024847807590

WHO-IPCS (World Health Organization/International Programme on Chemical Safety). 2002. *Global Assessment of the State-of-the-Science of Endocrine Disruptors*. Geneva, Switzerland: World Health Organization/International. http://www.who.int/ipcs/publications/new_issues/endocrine_disruptors/en/.

WHO-UNEP (World Health Organization/United Nations Environment Programme). 2013. *State of the Science of Endocrine Disrupting Chemicals - 2012*. Geneva, Switzerland: World Health Organization/United Nations Environment Programme. URL: http://www.who.int/ceh/publications/endocrine/en/.

Wiener, JB, Rogers, MD, Hammitt, JK, & Sand, PH. 2011. *The reality of precaution: Comparing risk regulation in the United States and Europe*. London: RFF Press.

WRR (The Netherlands Scientific Council for Government Policy). 2012. *Physical safety: A matter of balancing responsibilities*. The Hague: The Netherlands Scientific Council for Government Policy.

Zoeller, RT, Bergman, Å, Becher, G, Bjerregaard, P, Bornman, R, Brandt, I, Iguchi, T, et al. 2014. A Path Forward in the Debate over Health Impacts of Endocrine Disrupting Chemicals. *Environmental Health, 13* (1):1–11. doi:10.1186/1476-069X-13-118.

Summary

In the past, science and scientists were usually considered trustworthy and credible by default. However, in recent decades, a slow but steady shift to a sort of 'wariness by default' can be observed. Movements denouncing human vaccination programs ('antivaxxers') have gained global momentum and the roll out of 5G radiofrequency electromagnetic fields (RF EMFs) has been demonstrated against internationally, on the basis that current scientific insights on the safety of vaccinations or 5G RF EMFs are considered unreliable and untrustworthy. In the Netherlands, the science-based monitoring and modelling of nitrogen deposition in Natura 2000 sites, and derivation of (temporary) background values for PFAS in Dutch soil are the latest of a long list of issues where scientific methodologies, data or conclusions were openly disputed, dismissed or demonstrated against. Meanwhile, the availability of contrasting, or even mutually exclusive strings of scientific evidence allow stakeholders and policy makers to substantiate similarly contrasting policy preferences. In return, scientific advisory institutes, such as the Dutch National Institute for Public Health and the Environment (RIVM), are seeking ways to respond to these trends.

In this thesis, the issue of endocrine disrupting substances (EDSs) was selected as an example to study the reasoning and argumentation underlying international differences in stakeholders' regulatory preferences, in the context of ongoing scientific controversy. The main research question of this thesis is: what argumentation underlies the international differences in preferences of stakeholders towards the regulation of EDSs? **Chapter 1** described how EDSs became controversial, what we think are the dominant scientific controversies and why regulation of EDSs appears to remain a difficult endeavor. Or in short: what makes the issue of EDSs a highly suitable topic choice for this thesis. Also, several pertinent scientific theories and ideas about the role of scientific advisory institutes in contemporary society and policy are briefly introduced. The chapter concludes with describing a small exploratory study, where semi-structured interviews were used to assess why France specifically targeted thermal paper, as one of various applications of Bisphenol A (BPA), for drafting a proposal to restrict its use in Europe. On the basis of the results, we argue that (implicit) advocacy coalitions could be discerned in France; stakeholders from industry explicitly referred to one another to discern their specific standpoints and argumentation, and NGO's appeared to cooperate to inform the French population on the risks of exposure to EDSs and influence French policy makers.

Our research question can be approached from a wide variety of scientific directions, and a range of complementary answers may be possible, depending on the adopted disciplinary perspective. Accordingly, a variety of conceptual frameworks can be used to theoretically reflect on this matter, where each of these frameworks will provide a different piece of the puzzle. **Chapter 2** outlines the theoretical foundation for the present thesis, where eight conceptual frameworks were explored to theoretically reflect on the main research question. Specific attention went out to the role of (controversial) scientific knowledge in international differences in regulation of environmental health risk issues. The frameworks included in this

review are the Risk Assessment Paradigm, research into the roles of experts as policy advisors, the Psychometric Paradigm, the Cultural Theory of Risk, participatory approaches to risk assessment and risk management, the Advocacy Coalition Framework, the Social Amplification of Risk Framework, and Hofstede's Model of National Cultures. From the ideas and concepts offered by the eight frameworks, pertinent questions were derived to be used in support of our further empirical work, which was started in parallel to this theoretical review. The review led to the development of an overarching framework that illustrates the conceptual relationships between the eight frameworks. This overarching framework illustrates how competing advocacy coalitions can interact through the exchange of arguments, and how these arguments are based on the professional and personal knowledge, beliefs and (cultural) contexts of the coalition's members.

In **Chapter 3**, we zoomed in on the polarized scientific debate over EDSs and performed an argumentation analysis of two high-profile scientific papers (Lamb et al., 2014 and Bergman et al., 2015) that appeared to have conflicting perspectives on key issues in EDS science. We were interested in the underlying argumentation behind the conflicting positions, with respect to the interpretation of EDS science by these two groups of authors. We used Pragma-Dialectical Argumentation Theory (PDAT) to study whether the observed disagreements were instances of interpretative ambiguity over scientific knowledge or normative ambiguity over broader norms and values. The main result was that, while the analyzed discussion had the appearance of a purely scientific discussion, we indeed identified several instances of normative ambiguity. Two of these instances were studied in more detail using the typologies of Myths of Nature and Weiss' expert roles, to help underpin the normative elements at stake. The chapter concluded with arguing that, rather than attempting to settle such normative disagreements solely in the scientific sphere, broader stakeholder involvement should be sought.

In **Chapter 4,** our attention shifted to the debate about the EU's attempts to regulate EDSs**.** This chapter presents our analysis of the argumentation contained in responses to the EU's public consultation, concerning the impact assessment of four options to identify EDSs for regulatory concerns. The results of this consultation, summarized in EC (2015), show that different regulatory options are supported; some governmental entities (e.g. national governments) and stakeholder organizations (e.g. industry organizations) prefer an option including potency or even other considerations of hazard characterization, whereas others prefer an option that avoids these considerations and introduces several hazard-identification based weight-of-evidence categories. The main aim of this study was to analyze and compare the argumentation underlying different option preferences of several of these governmental entities and prominent stakeholder organizations, as stated in their responses to the consultation. We identified at least two overarching (implicit) advocacy coalitions, that held contrasting positions on five main themes, that include scientific, regulatory scientific, and regulatory aspects of options for regulatory criteria. We argued that the discussion on the EDS identification criteria requires a broad societal dialogue, in accordance with contemporary risk governance literature, where EDS science and regulation are explicitly discussed as interrelated themes.

Finally, **Chapter 5** provided an overall reflection on the work performed for the research project described in this thesis. We highlighted some links between our theoretical work and the empirical findings. Also, some suggestions on how to proceed in the debates on EDS science, policy and regulation were provided. Firstly, on the part of scientific experts, transparency about personal and professional values present in their research and advisory work is needed. Secondly, on the part of policy makers, awareness that disputes over broader norms and values should be settled in the political rather than scientific arena should be stimulate, even though the dispute appears to be purely scientific/technical on face value. And accordingly, that inclusive risk governance approaches are more suitable to deal with value-laden scientific issues, rather than unconditionally following the credo 'more research is needed'. Several theoretical frameworks presenting such inclusive approaches have been developed, and these could be field tested using the issue of EDSs or other environmental health risk issues. With regard to our uses of Pragma-Dialectical Argumentation Theory, we argue that, while the use of PDAT in this PhD research project may have been worthwhile, the accompanying analyses typically have a lead time of at least several months. This would make PDAT, at least as we used it, less suitable for smaller, less flexible or highly time-constrained projects. The development of tools for argumentation analysis that may be faster, easier and simpler to use, should be encouraged to facilitate the use of PDAT in such projects.

Samenvatting

In het verleden genoten wetenschap en wetenschappers een soort vanzelfsprekende betrouwbaarheid en geloofwaardigheid. In de afgelopen decennia is dit echter langzaam maar zeker omgeslagen naar scepsis. 'Antivaxxers' - groepen die ageren tegen grootschalige preventieve vaccinatiecampagnes - hebben mondiaal meer tractie gekregen. Op verschillende plekken in de wereld hebben demonstraties plaatsgevonden tegen de uitrol van 5G radiofrequente elektromagnetische velden (RF EMFs), met het in brand zetten van UMTS masten als (voorlopige) opzienbarende climax. Mede aan de basis hiervan ligt de overtuiging dat de huidige wetenschappelijke inzichten over de veiligheid van vaccinaties of 5G RF EMFs ongeloofwaardig zijn. In Nederland liggen de controversen rondom de monitoring en modellering van stikstofdepositie in Natura 2000 gebieden en rondom de afleiding van PFAS normen voor (te bebouwen) bodem nog vers in het geheugen. Deze controversen zijn de meest recente voorbeelden van een lange lijst momenten waarin onderzoeksmethodologiën, wetenschappelijke data of de daaruit getrokken conclusies openlijk in twijfel getrokken of verworpen werden, of waartegen zelfs gedemonstreerd werd. De beschikbaarheid van contrasterende, of zelfs elkaar uitsluitende wetenschappelijke resultaten, inzichten en argumenten stellen stakeholders en beleidsmakers in staat om eveneens contrasterende beleidsvoorkeuren te onderbouwen. Als gevolg hiervan zoeken wetenschappelijke instituten zoals het Rijksinstituut voor Volksgezondheid en Milieu (RIVM) naar manieren om met deze complexe maatschappelijke ontwikkelingen om te gaan.

In dit proefschrift is de casus 'hormoonverstorende stoffen' gekozen om redeneringen en argumentaties te bestuderen die ten grondslag liggen aan internationale verschillen in de regulatoire voorkeuren van stakeholders, in de context van vóórtdurende wetenschappelijk controverse. De hoofdonderzoeksvraag in dit proefschrift is dan ook: welke argumentaties liggen ten grondslag aan internationale verschillen in stakeholders' regulatoire voorkeuren voor hormoonverstorende stoffen? **Hoofdstuk 1** beschrijft hoe hormoonverstorende stoffen controversieel werden, wat wij denken dat de dominante wetenschappelijke controversen zijn en waarom wetgeving voor hormoonverstorende stoffen een ingewikkelde zaak blijkt te zijn. Of simpeler geformuleerd: wat maakt dat de hormoonverstorende stoffen casus zo geschikt is als onderwerp van dit proefschrift? Verder beschrijven we enkele bestaande wetenschappelijke theorieën en ideeën over de rol van wetenschappelijke adviesinstituten in de moderne samenleving en hedendaagse beleidsprocessen. Het hoofdstuk sluit af met een korte beschrijving van verkennend deelonderzoek. Daarin onderzochten we waarom de Franse overheid het gebruik van juíst thermisch papier, als één van de toepassingen van Bisfenol A, aanzienlijk wilde beperken in Europa door het opstellen van een restrictievoorstel. Op basis van de gevonden resultaten beargumenteren we dat (impliciete) 'advocacy coalitions' [een beleidskundig concept van Sabatier en Jenkins-Smith, 1993] geïdentificeerd kunnen worden in Frankrijk. In meerdere gevallen refereerden verschillende industrie-partijen namelijk naar elkaar om hun standpunten en onderliggende argumentaties gecoördineerd naar buiten te brengen. Daarnaast leken enkele NGO's samen te werken bij het informeren van de Franse

samenleving over de risico's van blootstelling aan hormoonverstorende stoffen, en bij het uitvoeren van beïnvloedingsactiviteiten gericht op Franse beleidsmakers.

Onze hoofdonderzoeksvraag – welke argumentaties liggen ten grondslag aan internationale verschillen in stakeholders' regulatoire voorkeuren voor hormoonverstorende stoffen? – kan vanuit verschillende wetenschappelijke richtingen benaderd worden. Dit leidt tot verschillende complementaire antwoorden en inzichten, want ieder wetenschappelijk vakgebied biedt immers een ander, aanvullend perspectief op de materie. Dat betekent dat verschillende 'conceptuele raamwerken' [theoretische modellen of raamwerken waarin wetenschappelijke concepten en ideeën schematisch worden weergegeven] gebruikt kunnen worden om vanuit theoretisch perspectief te reflecteren op onze onderzoeksvraag. Daarbij zal ieder conceptueel raamwerk een ander stukje van de puzzel bieden. **Hoofdstuk 2** beschrijft een theoretische review waarin acht conceptuele raamwerken de revue passeren. Dit hoofdstuk vormt daarmee het theoretisch fundament van dit proefschrift. We waren daarbij extra geïnteresseerd in de rol van (controversiële) wetenschappelijke kennis in internationale verschillen in wetgeving op het gebied van milieugezondheidsvraagstukken in brede zin. De volgende raamwerken zijn geïncludeerd in de review: 1) het Risk Assessment Paradigm, 2) de rol van experts als beleidsadviseur, 3) het Psychometric Paradigm, 4) de Cultural Theory of Risk, 5) participatieve methoden voor risicobeoordeling en -management, 6) het Advocacy Coalition Framework, 7) het Social Amplification of Risk Framework, en 8) Hofstede's Model of Nationale Cultures. Op basis van de ideeën en concepten van de acht raamwerken hebben we een lijst relevante vragen voor toekomstig empirisch onderzoek afgeleid. Dit empirisch werk was gelijktijdig gestart met deze theoretische review. Op basis van de review konden we tevens een eigen overkoepelend raamwerk ontwikkelen (figuur 3.12 op pagina 56 in dit proefschrift), waarin de onderlinge conceptuele relaties tussen de acht raamwerken geïllustreerd worden. In het kort laat ons overkoepelde raamwerk zien hoe concurrerende 'advocacy coalitions' kunnen (inter)acteren door het uitwisselen van argumentaties, en hoe deze argumentaties vervolgens gebaseerd zijn op professionele en persoonlijke kennis, overtuigingen en (culturele) contexten van de leden van de coalities.

In **hoofdstuk 3** staat het gepolariseerde wetenschappelijke debat over hormoonverstorende stoffen centraal. We hebben twee belanghebbende wetenschappelijke papers (Lamb et al., 2014 en Bergman et al., 2015) onderworpen aan een argumentatie-analyse. Deze twee papers lijken namelijk tegenstrijdige opvattingen te hebben over belangrijke wetenschappelijke thema's ten aanzien van hormoonverstorende stoffen. Het ging ons daarbij specifiek om de onderliggende argumentaties, en we hebben daarom Pragma-Dialectische Argumentatietheorie (PDAT) toegepast om te ontdekken of de geobserveerde meningsverschillen voortkomen uit interpretatieve ambiguïteit over wetenschappelijke kennis of normatieve ambiguïteit over bredere normen en waarden. De belangrijkste bevinding was dat, hoewel het geanalyseerde debat ogenschijnlijk puur wetenschappelijk leek, we inderdaad enkele gevallen van normatieve ambiguïteit hebben kunnen identificeren. Twee van deze gevallen hebben we nader bestudeerd door gebruik te maken van de Myths of Nature en Weiss' expert rollen typologieën. Dit gaf extra inzicht in enkele specifieke betwiste normatieve elementen. Het hoofdstuk sluit af met de

stelling dat het, daar waar het gevallen van normatieve ambiguïteit betreft, verstandiger lijkt om deze met brede stakeholder betrokkenheid aan te pakken dan te blijven pogen deze in de wetenschappelijke arena op te lossen.

In **hoofdstuk 4** verschoof onze aandacht naar wetgeving over hormoonverstorende stoffen, of specifieker: naar het lopende debat over de pogingen van de EU om hormoonverstorende stoffen te reguleren. Dit hoofdstuk bevat daarom een analyse van de argumentaties aanwezig in de responses van belanghebbenden op een publieke consultatie van de EU. Deze consultatie was onderdeel van een 'impact assessment' van vier uiteenlopende opties voor regulatoire criteria om hormoonverstorende stoffen van 'regulatory concern' te identificeren. De resultaten van deze consultatie, samengevat in EC (2015), lieten al zien dat verschillende belanghebbenden verschillende opties steunen. Dat wil zeggen, sommige overheden en andere belanghebbenden (bijvoorbeeld industriekoepelorganisaties) stelden een voorkeur te hebben voor een optie waarin potency [maat voor de toxiciteit van een stof; hoeveel stof er nodig is om een bepaald effect te geven] of zelfs andere elementen van 'hazard characterization' geïncludeerd worden. Anderen stelden juist de voorkeur te hebben voor een optie waarin deze elementen vermeden worden, maar juist enkele 'weight-of-evidence' categorieën [hoeveelheid, robuustheid en eenduidigheid van de beschikbare data benodigd om een stof als potentiële, vermeende of daadwerkelijke hormoonverstorende stof te classificeren] worden geïntroduceerd op basis van 'hazard identification' overwegingen. Het hoofddoel van dit onderzoek was om de argumentaties onderliggend aan deze optievoorkeuren te analyseren en vergelijken, en daarbij de responses van enkele overheden en andere belanghebbenden te gebruiken. Het resultaat was dat we in ieder geval twee overkoepelende (impliciete) 'advocacy coalitions' hebben kunnen identificeren die tegengestelde standpunten hadden op vijf kernthema's, zoals wetenschappelijke, regulatoire-wetenschappelijke en regulatoire eigenschappen van de vier opties. Tot slot concludeerden we dat de discussie over EU wetgeving van hormoonverstorende stoffen een brede maatschappelijke dialoog zou kunnen gebruiken, waarbij geput wordt uit wetenschappelijke inzichten op het gebied van risico governance, en waar hormoonverstorende stoffen wetenschap en wetgeving expliciet als onderling verbonden werelden worden behandeld.

In **hoofdstuk 5** geven we tot slot een algemene reflectie op het onderzoek dat is beschreven in dit proefschrift. Om te beginnen hebben we enige verbindingen tussen ons theoretische en empirische werk beschreven. Daarnaast hebben we enkele suggesties gedaan over hoe een constructief vervolg gegeven zou kunnen worden aan de debatten over hormoonverstorende stoffen wetenschap, beleid en wetgeving. Ten eerste is het voor wetenschappelijke experts noodzakelijk om transparant te zijn over de persoonlijke en professionele waarden die inherent - maar vaak impliciet - aanwezig zijn in hun onderzoek en advieswerk. Ten tweede is het voor beleidsmakers belangrijk om zich bewust te zijn dat geschillen over bredere normen en waarden beter aangepakt kunnen worden in een politiek/bestuurskundige context dan in een wetenschappelijke, ook al heeft het de schijn van een puur wetenschappelijk of technisch geschil. En dat daarmee inclusieve risico governance aanpakken beter passend zijn voor waardegeladen wetenschappelijke controversen dan het onvoorwaardelijk volgen van het credo

'meer onderzoek is nodig'. Er zijn al verschillende theoretische raamwerken ontwikkeld waarin dergelijke inclusieve aanpakken centraal staan, en hun praktijkwaarde zou goed getest kunnen worden door deze toe te passen op de hormoonverstorende stoffen casus of andere milieugezondheidsvraagstukken. Tot slot hebben we nog enkele reflecties gedeeld over ons gebruik van Pragma-dialectische argumentatie theorie (PDAT), welke we hebben gebruikt voor onze argumentatie-analyses. We stellen dat gebruik van PDAT voor dit PhD project de moeite waard is geweest, maar dat het wellicht minder toegankelijk is voor projecten die kleiner of minder flexibel zijn of minder doorlooptijd hebben. Dat komt voornamelijk door de uitvoeringstijd van de analyses, die minstens meerdere maanden bedroeg. De ontwikkeling van snellere, makkelijkere en eenvoudigere methoden voor argumentatie-analyse zou PDAT dus breder toepasbaar kunnen maken, en dat moedigen wij van harte aan.

Dankwoord

En ineens is het klaar! Ik weet nog goed dat het startschot een eerste meeting met onze projectgroep was, begin april 2015, destijds nog in het Cohengebouw. Natuurlijk landden de vele ideeën van mijn promotoren en copromotoren nog niet echt, zoals dat gaat. Nu, halverwege 2021, ligt er toch maar mooi een proefschrift met een kop en een staart. Het werk aan dit proefschrift was razend interessant en vaak leuk, en soms behoorlijk pittig. Ik heb ongelofelijk veel mogen leren. Over de onderwerpen die in dit proefschrift voorbij komen, over het RIVM, over wetenschap in brede zin, en over mezelf. Ontzettend veel mensen hebben mij op hun eigen manier geholpen dit proefschrift te schrijven. Vaak heel bewust, soms ook onbewust, maar daarmee niet minder belangrijk. Hen wil ik allemaal oprecht bedanken voor hun betrokkenheid.

Aldert, je tomeloze enthousiasme heeft me altijd geïnspireerd, ook (juist…!) als het even tegenzat. Je bent een ontzettend betrokken promotor, en er was altijd even tijd voor een praatje. Vaak natuurlijk erg inhoudelijk, maar soms ook heel persoonlijk. Dat schept veel vertrouwen en respect. Ik denk dat iedere promovendus zou boffen met jou als eerste promotor.

Erik, ik heb me vaak afgevraagd wat je eigenlijk níét weet. Je hebt ontzettend brede interesses en enorme parate kennis. Je hebt je altijd gepositioneerd als tweede promotor, maar was toch heel betrokken bij mijn werk. Ik kan me nog goed herinneren hoe Irene, jij en ik een reeks vrij intense brainstormsessies hadden om het theoretisch paper (en het uiteindelijke overkoepelende raamwerk) vorm te geven. Als er iemand is die mijn ideeën en papers naar een hoger plan tilde, ben jij het wel.

Theo (V.), al vrij snel in het traject heb jij het projectleiderschapsstokje overgenomen. Dat bleek een gouden greep. In het begin van mijn 'RIVM-carrière' was je al even mijn leidinggevende bij VSP, dus ik wist al dat je zeer kundig bent, naast vriendelijk en rustig in de omgang. Maar daarnaast bleek je ook een zeer betrokken en hulpvaardige projectleider. Dat heeft mij erg geholpen, dank!

Irene, van alle (co)promotoren is jouw naam bij ons thuis het meest gevallen. Want hóé vaak is het wel niet voorgekomen dat ik Erica heb moeten bekennen dat ik (veel) later thuis zou zijn omdat ik in een uren durende - en altijd erg interessante! - verhandeling met 'Irene van het RIVM' beland was. Je zit altijd barstensvol kennis, boordevol ideeën, en dat met passie voor het vak. En ik kon altijd bij je terecht als er wat te spuien was. Ik vind dat wij heerlijk hebben kunnen samenwerken, top!

Betty, je hebt een ongelofelijk scherpe geest, bent zeer betrokken bij het hormoonverstorende stoffen dossier en je hebt je met veel energie ingezet voor TARGETS. We hebben behoorlijk wat pittige discussies gehad, waarbij we het niet altijd eens konden worden. Misschien heb ik

ook juist daarom zoveel van je geleerd; van jouw kundigheid, en omdat je me bij uitstek dwong om met een gedegen verdediging te komen als ik een standpunt innam. Daar ben ik je zeker dankbaar voor.

Holly en **Lana**, het was ontzettend leuk en eervol dat ik jullie heb mogen begeleiden tijdens jullie stage bij ons TARGETS project. Jullie hebben beiden hard gewerkt, en jullie inzet heeft mede aan de basis gestaan van twee van de drie papers. Vanzelfsprekend ben ik jullie daar dankbaar voor!

Ronald, toen ik ergens begin 2015 de knoop had doorgehakt om aan de slag te gaan als promovendus bij het RIVM, spraken we beiden de wens uit dat ik bij jou in de afdeling zou blijven. Jij wilde deze 'reis' met mij doormaken, en ik vond jouw coachende en stimulerende leiderschapsstijl erg prettig werken. Ik heb ervaren dat je kunt mee- en tegendenken als geen ander, en dat scherp taalgebruik en haarfijn formuleren voor mij soms misschien best wel een ding is, en tegelijkertijd wel een hele belangrijke vaardigheid. Dank voor alle adviezen!

Natuurlijk gaat er grote dank uit naar mijn twee paranimfen: **Inge** en **Yolanda**. Ontzettend leuk dat we dit als Amersfoorts clubje doen, en dat jullie ook nog eens twee straten van elkaar wonen. **Yolanda**, je hebt flink zitten zweten op de omslag van dit proefschrift, maar het resultaat mag er helemaal zijn! En **Inge**, al na ons eerste gesprek met z'n drieën had je elke mogelijke vraag over het hoe-en-wat van de promotiedag gesteld, dat werkt fantastisch! Super dat jullie mij zo helpen.

Ik heb zo'n zes jaar bij **centrum DMG** rondgelopen, zonder er ooit echt in dienst geweest te zijn. Toch heb ik me er ongelofelijk thuis gevoeld. Dat zegt toch wel iets over hoe prettig de collega's zijn! Alleen al de lol en gezelligheid met onze **DMG-band** (of 'Public Servants', zoals we ons verkochten bij de lunchpauzeconcerten en andere optredens), met als harde kern **Ilse G., Rik, Theo T. en Theo van A.,** en ook **Harm**, **Michiel Z.**, **Pieter** en zelfs bandvriend **Thomas**. En **Francien** als trouwe 'groupie'. Wat was het leuk om met jullie al die repetities en optredens te doen, tussen al het werk door. Ik denk nog vaak aan onze hele puike vertolking van 'Sultans of Swing'! **Jochem**, wij hebben een toffe PhD-tijd gehad samen! We begonnen ongeveer tegelijk als promovendus. Eerst borrelden en aten we wat, maar dat monde gauw uit in reguliere squash-dates, de ultieme afleiding van het werk. Ook nu je in Boston zit hebben we nog contact, heel leuk! Verder was het ook met de andere **DMG Young** collega's erg gezellig. Hoe kan het ook anders met ál die borrels, feestjes, dinertjes en heuze sportactiviteiten. Dus dank **Inge, Kim, Nina, Ilse, Maud, Thomas, Martina, Ingrid, Julia, Peter, Erik (D.) en alle andere DMG Youngsters**. En ook al was ik als 'IRA2-promovendus' bij de **RIC Junioren** een beetje een vreemde eend in de bijt, het was het erg gezellig en leerzaam (niet in de laatste plaats vanwege het vrijuit kunnen spuien en miepen over het wel en wee van het promovendus-zijn). Dank **Tom, Linda, Kim, Marion, Joram en anderen!**

Ik wil ook het **IRAS** hartelijk danken voor hoe welkom ik er altijd ben geweest. De PhD-days, BBQs, en andere activiteiten waren erg geslaagd.

Daarnaast gaat mijn oprechte dank ook uit naar alle collega's bij **BDR**. Want hoewel ons figuurlijke huwelijk niet zo lang was, heb ik wel alle ruimte en steun gevoeld om goed aan mijn proefschrift door te blijven werken. Voor mij een niet te onderschatten bijdrage. Jullie zijn fantastische collega's geweest! In het bijzonder geldt dat voor **Team Internationaal**, en natuurlijk **Carla** en **Eric**. Erg leuk dat het contact RIVM-aanstellingen overstijgt!

Ruud, we zijn al een poos geen huisgenoten meer, maar ik wil je toch nog even genoemd hebben. Je was altijd geïnteresseerd, en verbaasde je tegelijkertijd altijd over de doorlooptijd van mijn publicaties. Hoe vaak ik wel niet 'nog éventjes' heb geantwoord op je vraag wanneer mijn papers geaccepteerd zouden worden!

Hannie en **Peter**, ik heb het getroffen met jullie als schoonouders. Jullie waren altijd oprecht geïnteresseerd in mijn werk, en bleven actief doorvragen. Ik weet nog goed dat jullie vertelden dat mijn dementerende schoon-oma aan jullie had gevraagd waar mijn onderzoek ook alweer over ging. Omdat je (Hannie) het écht simpel wilde houden, zei je maar 'milieu'. Zo bondig kan het dus ook!

Natuurlijk was er gelukkig ook de nodige ontspanning! Dank aan **Peter, Henry, Joram, Jeroen, Rian, Julian en Wouter.** Technisch gezien is er al (vriendschappelijke) chemie sinds '06 (sorry, not sorry), maar als promovendus ging het snel op de leuk-om-te-meeten-maar-wel-om-22-uur-naar-huis wijze (die ik sowieso al wat aan de vroegere kant had). Dank voor de gezelligheid! Ook inmiddels zeker richting **Lotte**, **Florentine**, en **Elien**!

Voor muzikaal vertier, de nodige (muziek)weekendjes weg en algehele vriendschappelijkheden waren natuurlijk **Daan, Leon, Olav, Thomas, Xander** en **Yolanda** van de partij**.** Promoveren is natuurlijk leuk, als er maar regelmatig s'avonds muziek gemaakt kan worden, moeilijk gedaan kan worden over makkelijke muziek, of de laatste muzikale hebbedingen uitgebreid doorgenomen kunnen worden! En laat ik de soort van 'running gag' niet vergeten, dat ik standaard gevraagd werd, als er ergens ter wereld iets speelde: 'en wat vindt het RIVM daarvan…?'. Dank allen!

Mick, Nischal en **Zeir**, heel tof hoe wij al láng leuke dingen doen samen, maar ook goede gesprekken hebben, en jullie mij gesteund hebben als het promoveren en mijn carrierekeuzes mij aanvlogen. We kunnen lang en uitgebreid praten (oké, misschien ligt dat soms wel aan de .ppt's van mij, maar toch), over waar we persoonlijk en qua werk allemaal tegenaan lopen, en welke kant het op kan. Of het gaat gewoon over voetbal, bitcoins of mooie herinneringen. Dank voor de gezelligheid en steun! Dat geldt ook voor **Brechtje**, **Daphne** en **Nejla**.

Papa, er was niemand geïnteresseerder in mijn onderzoek en de vele details dan jij. Opgeteld hebben we het úren en úren en úren over mijn onderzoek gehad, waarbij je aandachtig luisterde en altijd goed doorvroeg. Prachtig als iemand net zo enthousiast is over waar ik mee bezig ben als ikzelf. En **mama**, jij bent er, zoals altijd al, voor me geweest om te helpen waar

nodig. Voor steun, opbeurende woorden en reflectie. Ik ben dankbaar dat er altijd een opgemaakt bed voor me klaarstond om even afstand van het werk te nemen. En dan natuurlijk **Sylke**, **Rogier**, **Esmé** en **Jack**. Wat is het gezellig om met jullie te zijn, om te spelen met het prachtige dametje en heertje of simpelweg bij te kletsen. Gewoon bij jullie thuis rondhangen was hét antwoord op promovendus-stress.

Lieve **Erica**. Sinds kort Erica Clahsen. Woorden schieten te kort voor wat je voor mij hebt betekend tijdens dit traject. Je memoreert zo nu en dan nog dat het voor jou toch aanlokkelijk was dat ik op mijn Tinder-profiel 'Promovendus @ RIVM' had staan. Ik vraag me af of je je wel besefte wat dat in de praktijk kon gaan betekenen ;-) De pieken bleken gelukkig hoog, de dalen soms diep. Maar je was er, in voor- en tegenspoed! Topper!

En daarmee is mijn proefschrift afgerond. Aan iedereen die ik nog vergeten ben: hartelijk dank!

About the author

Sander Carl Stijn Clahsen was born September 28th 1988 in Enschede, the Netherlands. He graduated from high school (Het Stedelijk Lyceum, location 'Zuid') in 2006. He started out as a 'beta' by obtaining a Bachelor of Science degree in Chemical Engineering from the University of Twente in 2011. However, near the end of this study, he realized he was more into exploring the societal impact of science and technology, rather than becoming a practicing engineer himself. So he started a master in Science Communication, also at the University of Twente, which was completed by performing an internship at the RIVM (Dutch National Institute for Public Health and the Environment) in Bilthoven. After he received his Master of Science degree in 2013, he remained at RIVM and kept working there for about seven years, in different professional roles: he was scientific employee for one-and-a-half years, until he started working on this PhD thesis for four years (in collaboration with IRAS/Utrecht University), and then became policy advisor on international affairs for one year. In June 2021, he made a switch to a position in higher education teaching, by starting as a lecturer in Public Management (in Dutch: Bestuurskunde) at The Hague University of Applied Sciences (in Dutch: de Haagse Hogeschool).

Sander is now well versed in the complex dynamics occurring in the science–policy-society triangle, particularly where environmental health issues are concerned and scientific knowledge often is complex, uncertain and contested.

List of publications

PUBLICATIONS IN THIS THESIS

Sander C.S. Clahsen, Irene van Kamp, Betty C. Hakkert, Theo G. Vermeire, Aldert H. Piersma and Erik Lebret. 2019. Why do countries regulate environmental health risks differently? A theoretical perspective. *Risk Analysis 39 (2)*: 439-461.

Sander C.S. Clahsen, Holly S. van Klaveren, Theo G. Vermeire, Irene van Kamp, Bart Garssen, Aldert H. Piersma and Erik Lebret. 2020. Understanding conflicting views of endocrine disruptor experts: a pilot study using argumentation analysis. *Journal of Risk Research 23 (1)*: 62-80.

Sander C.S. Clahsen, Lana Moss, Irene van Kamp, Theo G. Vermeire, Bart J. Garssen, Aldert H. Piersma, Erik Lebret. 2020. Analysis of different preferences for the EU's regulatory options for endocrine disruptor identification criteria using argumentation theory. *Science of the Total Environment 740*.

OTHER PUBLICATIONS

Rik P. Bogers, Anne van Gils, Sander C.S. Clahsen, Wendy Vercruijsse, Irene van Kamp, Christos Baliatsas, Judith G.M. Rosmalen, John F.B. Bolte. 2018. Individual variation in temporal relationships between exposure to radiofrequency electromagnetic fields and non-specific physical symptoms: A new approach in studying 'electrosensitivity'. *Environment International 121*: 297-307.

John F.B. Bolte, Sander C.S. Clahsen, Wendy Vercruijsse, Jan H. Houtveen, Maarten A. Schipper, Irene van Kamp, Rik Bogers. 2019. Ecological momentary assessment study of exposure to radiofrequency electromagnetic fields and non-specific physical symptoms withself-declared electrosensitives. *Environment International 131*.