CHARACTERIZING ARGUMENTS ABOUT ENDOCRINE DISRUPTORS AND HUMAN HEALTH

by

Georgia Green

B.Sc., Imperial College London, 2019

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(RESOURCES, ENVIRONMENT AND SUSTAINABILITY)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

January 2023

© Georgia Green, 2023

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, the thesis entitled:

CHARACTERIZING ARGUMENTS ABOUT ENDOCRINE DISRUPTORS AND HUMAN HEALTH

submitted by Georgia Green in partial fulfilment of the requirements for

the degree of Master of Science

in Resources, Environment and Sustainability

Examining Committee:

Gunilla Öberg, Professor, Insitute of Resources, Environment and Sustainability, University of British Columbia
Supervisor

Annegaaike Leopold, Adjunct Professor, Insitute of Resources Environment and Sustainability University of British Columbia
Co-supervisor

Kevin Elliott, Professor, Department of Fisheries and Wildlife, Michigan State University
Additional Examiner

Abstract

Scientific controversies surrounding complex questions of societal concern are at times invoked as reasons to delay policy action. Previous research suggests there are different and potentially irreconcilable ways of knowing in the sciences.

The aim of the present study is to explore if distinct arguments can a) be detected in scientific literature about endocrine disruptors in relation to human health and, b) are rooted in different perceptions of the types of knowledge considered reliable, relevant, and valuable for informing chemicals risk assessment.

The analysis led to the identification of a series of arguments that differed in their starting points, the type of inferences they claim can be drawn from certain evidence, the amount of evidence needed to draw a conclusion and the research questions that ought to be answered. The arguments placed different weights on methods commonly used for evaluating the link between EDs and human health, including mechanistic studies, epidemiological studies and animal studies.

The present study suggests that policy relevant debates are happening in ED scientific literature. The findings also indicate that the operation of chemical regulatory systems to some extent dictate the knowledge being produced in the literature because findings are being developed with policy applications in mind, although this could be a pragmatic approach for scientists to ensure their work is compatible for the existing policy operations.

Lay Summary

Chemicals that pose a risk to human and environmental health are increasingly present in the natural environment. Some of these chemical pollutants are known as endocrine disruptors that can adversely interfere with hormone function in humans. Due to the complexities of endocrine disruptor action, there is ongoing scientific disagreement about the link between these endocrine disrupting chemicals and human health. Previous research has suggested that there are different perceptions among scientists of how to produce reliable and useful scientific knowledge about chemical risk. Bibliometrics and analysis of scientific literature led to the characterization of a series of distinct arguments that differ in the way they value different methods, including animal studies and epidemiological studies, and advocate for different policies.

Preface

Identification and design of the project was conducted by Georgia Green with guidance from Gunilla Öberg and Annegaaike Leopold. The literature review, research and analysis was conducted by Georgia Green with guidance from Gunilla Öberg and Annegaaike Leopold.

This research was not subjected to an ethics review as it was not required.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AEA Analagous Effect Argument

CA Contextuality Argument

CHA Collective Harm Argument

DEA Difficulty in Extrapolation Argument

DECA Dual Evidence Causality Argument

EChA Epidemoiological Characterization Argument

ED Endocrine Disruptor

EHA Epidemoiological Harm Argument

ICA Individual Chemical Argument

ITA Inherent Toxicity Argument

LAA Lack of Association Argument

MCA Mechanistic Causality Argument

MEA Mechanistic Extrapolation Argument

RMA Reductive Mechanistic Argument

WOR Way of Reasoning

WOE Weight of Evidence

Acknowledgements

I would like to thank my supervisors Gunilla Oberg and Annegaaike Leopold. I am grateful for your support, enthusiasm and insights.

I would like to thank the Egesta Lab for all their help, especially Jerry Achar and Jack Durant whose contributions strengthened this work.

Thank you to my friends and family for their enduring support. Special thanks to Emma, Grace, Trisa, Mauricio and May.

I am grateful for the funding I received through the Walter C. Koerner Fellowship and to the Centre for Accessibility.

Dedication

To Ella Viale-Sole and Eunice Moon, as a thank you for your kindness, laughter, and unwavering support.

Chapter 1: Introduction

1.1 Endocrine disruptors (EDs)

Chemicals are an integral part of modern life. They are used in many consumer goods, in food packaging, clothing, in medication, cosmetics, in pesticides and plastics. When using products containing chemicals, humans can be exposed to these chemicals. After or during use, such as pesticide application or the excretion of pharmaceuticals, these chemicals can enter the environment. Chemicals can pollute water, soil and air, and can be consumed by animals. Humans can also be exposed to environmental pollutants through skin contact as well as through mediums such as drinking water and food.

Certain chemicals used in various products are endocrine disruptors (EDs). Endocrine disruptors are chemicals that can interfere with the action of hormones and cause adverse effects. WHO and UNEP use the definition of:

> *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. (United Nations Environment Programme & World Health Organization, 2013, p.4)*

Hormones regulate many of the body's processes, including metabolism and the menstrual cycle and have important roles in growth, development, and reproduction. Therefore, disruption of such pathways can result in adverse health outcomes.

Endocrine disruptors can disrupt hormone pathways through various mechanisms, for example, sometimes their structural similarities to natural hormones allow them to bind to the body's hormone receptors. As a result, normal hormone function can be blocked or otherwise influenced. Certain EDs can also bind with enzymes involved in hormone regulation to prevent them from acting. There are many different ways in which endocrine disruptors can impact the hormonal processes in the body.

Studies have also indicated that EDs can interfere with gene transcription (Amaro et al., 2014; Klinge, 2015) and alter the expression of genes (Rattan & Flaws, 2019).

Many widely used chemicals are EDs. For instance, compounds that are used in birth control pills, such as ethinylestradiol, are designed to mimic the hormone estrogen. Such chemicals have been found to be present in sewage and the environment (Adeel et al., 2017). Exposure to artificial estrogens has been found to correlate with increased incidence of breast and prostate cancer (Moore et al., 2016; Nelles et al., 2011).

Rachel Carson's popular science book *Silent Spring* is credited for drawing both public and scientific attention to the impacts of chemical pesticides on environmental and human health. Studies of the health effects of EDs have grown progressively since the 1991 Wingspread conference spearheaded by Theo Colborn and others and written about by Krimsky in their book *Hormonal Chaos* (Krimsky, 2000). Already in the 1950s and 60s, research suggested that many compounds introduced into the environment by human activity are capable of disrupting the endocrine system of animals, including wildlife and humans.

All in all, it is feared that endocrine disruption can have profound impacts because of the crucial role hormones play in controlling development. Research on human health effects of EDs is, however, characterized by a high level of uncertainty, due to the complexities associated with studying EDs. For example, firstly, humans can be exposed to multiple chemicals through multiple different pathways simultaneously. Secondly, exposure can happen over long periods of time, with exposure having different consequences depending on life stage (e.g. pre-natal, childhood, adolescent; (Meeker, 2012; Roy, 2009; Wen et al., 2017).

A basic principle of toxicology is that the dose makes the poison; the higher the exposure dose, the more severe the effect (credited to Paracelsus). This holds true for many chemicals. However, the action of EDs challenge this principle. Firstly, there is evidence that low-dose exposure over time can result in adverse effects (Lang, 2008; Patterson et al., 2015) (Patterson et al., 2015). There is also evidence suggesting short-term exposure to EDs during sensitive development phases can also result in adverse effects (Saeedi Saravi & Dehpour, 2016; Signorile et al., 2010). Studies of the health effects caused by EDs are further complicated by the fact that the processes can be mechanistically complex. There are studies that show that some EDs induce non-monotonic dose response curves, meaning that the severity of an adverse effect does not increase proportionally with a dose (Andrade et al., 2006; Vandenberg, 2014), or that some effects are only observed at low doses but not high doses (Angle et al., 2013).

Other studies show that some EDs can activate alternate pathways or initiate different effects depending on the exposure dose (Alonso-Magdalena et al., 2012; Taxvig et al., 2007). It is suggested that this could be because these EDs target multiple receptors (MacKay & Abizaid, 2018). There is also evidence that exposure to multiple chemicals can activate multiple pathways simultaneously, with each resulting in different effects (Goodson et al., 2015). Many studies are conducted on animals, such as rats, or at a sub-organism level (e.g., organs, cells, enzymes). Because of the biological differences between these species and humans, it can be difficult to use these studies to make inferences about human health. Due to these and other uncertainties, it is difficult to provide a solid picture of ED health effects.

1.2 ED, human health and scientific disagreement

The large complexities surrounding EDs have resulted in scientific disagreement about how to best assess human health effects from EDs. These debates have been analyzed in a handful of papers: (Autrup et al., 2015; Clahsen et al., 2020; Durant, 2021; Zoeller et al., 2014). Below, these are discussed in chronological order.

Zoeller et al (2014) argue that there are differences in interpretation for several common terms used in endocrine disruptor literature. For example, what constitutes an adverse effect is, according to Zoeller, debated with Nohynek et al., (2013) describing it as toxicity, meaning a pathology or functional impairment while the International Program on Chemical Safety (IPCS) (2004) describe it as a "change in morphology, physiology, growth, development, reproduction or life span of an organism, system, or (sub) population that results in an impairment of functional capacity." (p. 3, cited in Zoeller). Zoeller also argues that there are conceptual differences in how the endocrine system is understood, with some framing it as a means of homeostatic control and others as an important factor in development.

Autrup et al. (2015) present an analysis on what they term "the present debate on chemicals with hormonal activity." They find a key disagreement in the endocrine disruptor field to be between researchers that see no reason why traditional risk characterization approaches cannot be applied to endocrine disruptors, while other researchers argue that the unique toxicology of endocrine disruptors should result in them being treated differently, with the authors placing themselves solidly on the former side of the debate:

> *In our opinion, chemicals with hormonal activity can be subjected to the well-evaluated health risk characterization approach used for many years including adverse outcome pathways. Many of the points arguing for a specific approach for risk characterization of chemicals with hormonal activity are based on highly speculative conclusions. (Autrup et al., 2015, p. 11)*

As this paper places itself on one side of a debate it cannot be considered an unbiased characterization of the ED knowledge controversy.

Clahsen et al. (2020) applied argumentation analysis to two papers that they assessed as presenting conflicting views on uncertainty and dissent on norms and values. They found differences in argumentation around the degree of controversy within the ED scientific debate, the utility of weight of evidence approaches, how causality is established, how the endocrine system is framed, and whether or not a particular state of science report can be used to support evidence-based decisions. Their findings indicate that expert disagreement in the ED field is not purely about scientific knowledge itself; but also, about how this knowledge should be produced, perceived and valued. For example, Bergman et al., (2015) argue that Lamb et al. (2014), both cited in Clahsen, had ulterior motives for critiquing the UNEP-WHO report (United Nations Environment Programme & World Health Organization, 2013); they were sponsored by the chemical industry and wanted to generate uncertainty on the health impacts of endocrine disruptors.

In his master's thesis, Durant (2021) took a brokered dialogue approach (Lamb et al., 2014) to interviewing two experts that had been identified as representing a more risk-based and a more hazard-based way of thinking, respectively, based on their published papers. Durant found that the experts considered certain types of uncertainties more critical than others (methodological vs epistemic) and that they viewed the uncertainties as originating from different sources (the research itself, vs wider factors influencing the research system). Although Durant found that while the experts did have some common ground in advocating for diversity, data transparency and interdisciplinary research, they disagreed on topics such as extrapolation from animal studies, industry research and research required to reduce uncertainty.

The Zoeller and Autrup papers assume that these debates, or scientific controversies are due to the uncertainty in the field. They therefore propose that these uncertainties can only be resolved in a certain way; through more research from specific fields (Autrup), or by developing unbiased criteria for systematically reviewing ED research (Zoeller).

> *We strongly support a rational debate based on factual data and established principles of science and specifically of pharmacology and toxicology. Such principles remain relevant to assess potential health impacts of chemical exposures, and only the application of such principles will permit a path forward in discussing the potential health impacts of 'endocrine disruptors'. (Autrup et al., 2015, p.13)*
>
> *It is important to develop transparent, consistent and unbiased criteria for the systematic review of EDCs. (Zoeller et al., 2014, p.10)*

Clahsen et al. take a different stance and claim that these disagreements cannot be resolved without acknowledging the value judgements that underlie them.

> *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. (Clahsen et al., 2020, p. 78)*

Literature on scientific controversies has also suggested that further research does not always resolve controversies, particularly in policy relevant fields. Ziman (2000), for example, argues that 'science has no official machinery for closing a controversy.' In other words, as there is no commonly agreed definition of what constitutes sufficient evidence, consequently there is no commonly agreed upon marker for when enough research has been conducted to resolve a controversy. Examples of this are chronicled in the book Golem Science by Collins and Pinch (2001). One of their examples is the debate in physics about cold fusion: In order to resolve the controversy of whether experiments demonstrated cold fusion, various scientists repeated experiments and undertook new ones with varying results. The additional experiments did not lead one side of the controversy to concede to the other because each side continued to argue that the new results supported their conclusion, or that the results were still ambiguous.

It is well documented that scientific uncertainty can be used to delay policy action. Campbell (1985) describes how experts can disagree with a policy action by claiming there is too much uncertainty. Therefore, if experts want to contest a policy action, they can use the continuing uncertainty as justification against taking said action. In *Doubt is Their Product* (2008) Michaels provides several examples, including how the tobacco industry hired scientists to magnify and manufacture uncertainties in studies linking health effects to tobacco usage. Similarly, in *Merchants of Doubt* (2011) Oreskes and Conway argue that 'keeping the controversy alive' was used as a tactic to delay policy intervention through magnifying and manufacturing scientific uncertainty after a consensus was reached on various issues from DDT to global warming. In *The Fight Against Doubt* (de Melo-Martín & Intemann, 2018), the authors coin the term normatively inappropriate dissent (NID) for challenging scientific findings on the basis of ulterior motives. A cited example is that of the oil and gas companies contesting the evidence for climate change to protect the value of their companies.

The risks tied to NID are particularly important when the results of scientific controversies have implications for society. The link between EDs and human health effects is a highly policy relevant field, since this research informs chemicals management practices. Policy recommendations for chemicals management are commonly made in scientific literature (Cousins et al., 2016). A considerable portion of the literature published in academic journals is, however, not used for policy-purposes, in part because many academic studies are not conducted according to Good Laboratory Practice (Ingre-Kahns 2019; Agerstrand 2017). In Canada, as in most other jurisdictions, the authorities use a combination of data provided by industry, scientific literature and expert judgment to assess the risk of chemicals (Barton-Maclaren et al., 2022).

Various stakeholders including the chemical industry, those affected by chemical pollution and those that use chemicals and are exposed to them (*e.g.*, farmers, consumers) take an interest in chemicals management policy. There is evidence that, historically, the chemicals industry has attempted to supress or discredit research linking chemicals to human health effects, for example by contesting the claims made in Carson's *Silent Spring* (Souder, 2012). Based on a review of policy actions for EDs, Aho (2017) argues that industry has countered the regulation of EDs. Karlsson (2019) argues that, similar to climate denial, fake experts are employed for 'chemicals denial' by questioning the causal links between chemical exposure and health effects. In less extreme examples, there are documented examples of stakeholders having different views on how the benefits and risks of chemicals should be weighed against one another. For example, in their book *Risk, Science and Politics* Harrison & Hoberg, (1994) describe several case studies that illustrate how conflicting societal values have led to different regulatory strategies for chemicals. More recently, stakeholders have raised awareness of how more vulnerable communities are more exposed to chemical pollution, they therefore advocate for more equitable chemical management policies (Ford & Scott, 2017; Takaoka et al., 2014).

In summary, similar to other complex policy-relevant areas where science is contested, there are indications that knowledge controversies in the ED field are not only related to scientific facts, but also to societal values.

Earlier research in this area suggests that the common expectation that more research will eventually lead to scientific consensus disregards that value-judgements underlie the science. By disregarding these value judgements, scientists may end up on different sides of a debate, without understanding why this is, which in turn hampers policy development. McIlroy-Young et al. (2021) undertook narrative analysis of focus groups with scientists who were assessed as belonging to different sides of the ED debate. They found that the motives assigned by the groups to themselves and the other group differed, with each group casting the other group as a small group of dissenters challenging the consensus of the majority. The authors suggest that the expectation of consensus can lead to a situation of manufactured consensus, where a consensus is reached by excluding relevant perspectives and hence it is not a true consensus. This is a concern if it is a common occurrence in the process of chemical regulation, as potentially relevant perspectives would then be unduly excluded from policy processes.

The studies by McIlroy-Young et al. (2021) and Clahsen et al. (2020) suggest that scientific disagreement about how to best evaluate the potential harm posed by endocrine disruptors goes beyond the uncertainties in the field, and might be rooted in fundamental disagreements about the value of different types of knowledge. For the purposes of this study, the underlying logic of arguments posed in relation to EDs will be termed 'Ways of Reasoning'. The aim of the present study is to identify and characterize distinct arguments present in the peer reviewed ED literature and contribute to discussions about how diverging preferences given to different types of research might relate to epistemic disagreements, which, in turn, is assumed to facilitate the use of science for policy-purposes.

Chapter 2: Theoretical background

Various scholars have explored the process of knowledge creation and developed different ideas that provide insight into different perceptions of what constitutes valid knowledge. Below, I provide a summary of the scholars that have influenced my work.

2.1 Ludvik Flek

In 1936, Ludvik Flek, an Austrian physician who was active at the turn of the 19^{th} century, wrote a book about the process of knowledge creation. Flek (1936/1970/2008) used the history of syphilis as a case study to illustrate how scientific facts come to be. Specifically, he tracks how syphilis was understood to be a disease and how the Wassermann reaction was understood to be a reliable way to test for syphilis. Flek argues that these two facts were created, not discovered, because scientific facts are in part socially constructed. This is because he sees the process of cognition that leads to the genesis of a scientific fact as being somewhat socially conditioned. He claims that all cognition and perception happen through what he calls "thought styles". Thought styles are a way of acknowledging that the way in which people see reality is constrained. Under a thought style only certain questions are considered, and only certain answers are permissible.

Flek argues that many scientific advancements are seen as irrefutable facts because the scientific process is often framed through the lens of existing thought styles. For example, he points out that research is often presented as progressing linearly to meet a particular goal, rather than reflecting the sometimes-roundabout scientific process and ignoring the impact of social climate and thought styles on the process of knowledge creation. He also describes what he terms journal science as a key mechanism through which facts appear to become indisputable. Flek describes journal science as making a plea for findings to be collectively accepted as facts, so journal science is a way of legitimizing knowledge in a certain socially acceptable way. The act of findings being communicated and transmitted to a wider audience in this way is what makes them appear to be indisputable. Flek suggests that journal science is therefore a poor indication of actual scientific practice, suggesting that the process of knowledge legitimization can change the knowledge itself.

Flek proposes that if cognition or knowledge creation is a collective activity, then certain groups will share certain ways of thinking that may fundamentally differ from other groups' ways of thinking. He therefore introduces the idea of thought collectives, the entity at the social level through which thought styles are expressed and maintained. Thought collectives are a community of people mutually exchanging ideas or maintaining intellectual interaction. Members of a thought collective will have a prevailing thought style and adopt certain ways of perceiving and thinking. The assumptions held by a thought collective can potentially impact how their members participates in the knowledge creation process. If one thought collective dominates it could potentially impact how society as a whole participates in the knowledge creation process.

Flek also argues that there are two components of knowledge: active and passive. Passive knowledge is reality independent whilst active knowledge rests upon the certain dispositions, goals and assumptions adopted by the thought collective. The issue is that the two components of knowledge can rarely be distinguished from each other, making it difficult to identify the impact of thought collectives on the knowledge creation process. This could also make it difficult for people to realize if they are in a thought collective or to characterize and make explicit the passive assumptions of thought collectives. Importantly, it also raises the question of whether different thought collectives can communicate with each other. Flek's work provided a grounding for later scholars who also took a process like approach to understanding how different ways of knowing emerge. He is still influential today for scholars who investigate why different groups of people think very differently and fail to communicate with or understand one another.

2.2 Thomas Kuhn

Thomas Kuhn was an American philosopher of science. He completed a bachelors, masters, and PhD in physics before later switching to study the history of science. In 1962, he published his influential work *The Structure of Scientific Revolutions*. He was strongly influenced by Flek, to the extent that he ensured Flek's book was translated into English. Both scholars take a process-based perspective to investigate scientific knowledge creation. *The Structure of Scientific Revolutions* describes the phenomenon of paradigms. According to Kuhn, in mature science there are switches between what he terms normal science and revolutions. In normal science, the key theories, instruments, methods, and assumptions are kept the same, which he terms a disciplinary matrix. During periods of normal science there is a period of cumulative knowledge increase, but certain questions are left unanswered or unconsidered, of which some deal with fundamental issues. After a certain time, the number of anomalous results increase, that are not explainable under the current paradigm. This leads to period of exemplary science, where the disciplinary matrix undergoes serious revisions. The result is a paradigm shift, which can introduce new problems for scientists to work on, suggest new approaches to solving these problems and creates a new standard by which proposed solutions can be measured.

Paradigms appear to be similar to thought styles in the sense that they contain certain assumptions that dictate knowledge creation which can limit what questions can be asked under them. Both Kuhn and Flek appear to tie paradigms and thought collectives closely to the history and progress of science, and do not consider whether multiple thought collectives or paradigms can exist simultaneously. Kuhn does, however, describe different scientific paradigms as being incommensurable. This means that science conducted under different paradigms cannot be comparable to each other as there is no common measure for assessing the merit of different scientific theories. This appears to be the first notion of acknowledging that certain ways of scientific thinking can be incompatible with others.

2.3 Ian Hacking

Ian Hacking is a contemporary Canadian philosopher of science who was heavily influenced by Kuhn. After obtaining a PhD in 1962, Hacking held a variety of research positions including at the Centre for Interdisciplinary Research and the University of Toronto. In 1975 Hacking published *Why Does Language Matter to Philosophy?* and *The Emergence of Probability*. He was influenced by the work of Michel Foucalt, Thomas Kuhn, Paul Feyerabend and Imre Lakatos. Hacking does not deal with the process of knowledge creation directly but explores the different ways in which scientific truths can be expressed and validated. His work therefore has relevance for understanding why different scientists, particularly those from different disciplines and time periods, can have different perceptions of what counts as valid knowledge. In his 2012 book *Language, Truth and Reason*, Hacking is drawing heavily on the work of Kuhn when describing what he terms “styles of reasoning” (Hacking, 2012)

Hacking argues that there are multiple styles of reasoning, which, according to him, are different ways of reasoning scientific truth that have been observable throughout history. Although Hacking does not explicitly define what is meant by styles of reasoning, various scholars have interpreted them as different ways of expressing propositions that can be proven true or false. Therefore, different statements can be proven true or false under different styles. For example, the question statement ‘does chemical x cause disease y?’ could potentially be evaluated through a statistical style (e.g. an epidemiological study) or a hypothetical comparative model (e.g., bioinformatics.)

The styles of reasoning, as adapted from Sciortino (2017) are described below:

(a) The simple method of postulation exemplified by the Greek mathematical sciences.

(b) The deployment of experiment both to control postulation and to explore by observation and measurement.

(c) Hypothetical construction of comparative models.

(d) Ordering of variety by comparison and taxonomy.

(e) Statistical analysis of regularities of populations

f) Laboratory, artificially creating conditions for experiments that cannot be conducted in the natural world.

According to Hacking, advances in technology or the discovery of new methods can give rise to either new styles or increase the prominence of one style over the other within a discipline. Hacking himself states that these styles are not comprehensive, because there is always the potential for new styles to evolve and other styles may already exist but not have been characterized by him. Hacking also only considers styles of reasoning in the western science tradition, meaning that it is possible that other styles could exist outside this paradigm.

2.4 Hugh Lacey

Hugh Lacey is a contemporary, now retired Australian philosopher, focusing on values influencing scientific knowledge production. His research positions include Research Fellow and Coordinator of the Working Group in Agroecology and Group of Studies in Philosophy, History and Sociology of Science and Technology. His research focuses on the intersection between nature, science and values and includes topics such as agriculture and food systems, science, and technology. His work appears to align with Hacking's perspective that there can be different ways of reasoning scientific truth, but his work investigates the process of knowledge creation rather than the output. Lacey looks into what he terms 'methodological strategies' in scientific research, referring to the choices made throughout the research process. Lacey argues that different research projects have different methodological strategies. He says that these strategies act to constrain the theories and hypothesis that can be candidates for investigations, and by doing so constrains the methods, models and possibilities that can be considered. This also selects the type of empirical data that is required, and what categories can be used to interpret the data.

From this it can be inferred that Hacking's styles of reasoning and Lacey's methodological strategies can themselves be thought of as different epistemic perspectives, as they imply that different scientific questions are sometimes only investigated in different ways that can constrain how their scientific truth is evaluated. Both Hackings' styles of reasoning and Lacey's methodological strategies can lead to the exclusion of other possibilities. Hacking and Lacey's work also links back to Kuhn's work on incommensurability, by implying that different styles or methodological strategies cannot be directly comparable with each other.

2.5 Karin Knorr-Cetina

Karin Knorr-Cetina is a contemporary Austrian sociologist known for her work in epistemology and social construction. Knorr-Cetina achieved a PhD in cultural anthropology and sociology from the University of Vienna. Her work moved towards contemporary studies of science, publishing her first work on this subject in 1976, *Determinants and Controls of Scientific Development*. After moving to the University of Berkeley, she continued her empirical research on science publishing the *Manufacture of Knowledge: An Essay on the Constructivist and Contextual Nature of Science*. Her book, *Epistemic Cultures: How the Sciences Make Knowledge*, was published in 1999.

Whilst Flek and Kuhn take a more historical and process-based approach to understanding how scientific thinking chances over time, Knorr-Cetina instead focuses on the day-to-day activities of scientists in two different scientific disciplines to attempt to understand how the different disciplines create, validify and legitimize knowledge. In *Epistemic Cultures*, Knorr Cetina compares the process of knowledge creation in two different scientific disciplines: particle physics and molecular biology (Knorr-Cetina, 1999). She argues that each discipline has their different epistemic strategies, organizational structures, and behavioral norms.

In her studies of particle physics, Knorr-Cetina notes that scientists use a variety of detectors to help them characterize particles that are not directly observable. The prevailing assumption is that empirical data cannot be understood if the process of measurement and observation is not also understood. Readings on their own mean nothing if it is not understood how the detector works. Negative knowledge, including background noise, uncertainties, ambiguities, or problems with the instruments are not just seen as experimental limitations, but they are considered the first steps of future study. The eventual aim of the discipline is to advance understanding by pushing the limits of knowledge; what Knorr-Cetina terms liminal knowledge.

Knorr Cetina explores what she terms ‘epistemic strategies,’ meaning the approaches used to ensure and justify the validity of a result. A strategy in molecular physics is to frame the different ways in which objects and information relate to each other. For example, sometimes experimental components are designed to compete with one another to find the most promising solution, sometimes they are complementary, sometimes they are used to verify results or to eliminate fake signals. This framing also extends to whole experiments, with sister experiments being a key strategy. In particle physics, sister experiments run in parallel, sharing the same end goal but differing in their approach. They are not replications as the experiments therefore maintain a degree of independence and limited communication with each other and are seen as an important experimental strategy to ensure validity in particle physics.

Alternative epistemic strategies are found in molecular biology. In contrast to physics, interactions with natural phenomena in the laboratory is what defines this epistemic approach. Whilst physics attempts to indirectly construct reality, molecular biology shows preferences for experimental evidence that does not rely on speculative interpretation. To ensure correct interpretation, scientists are expected to have an experimental register, or an awareness of how their own ways of working and perceptions could influence their results. Manual and hands-on lab experience is also seen as essential in interpreting results, so much so that even lab supervisors like to see the results themselves. Knorr-Cetina finds that for molecular biology, the main strategy is the study of certain objects in a laboratory setting, such as cells or mice. These objects are transformed in a certain way, such as genetic modification or tumor implantation. These transformations are often completed by following certain step by step experimental protocols, that are often traded between labs. When these protocols lead to ambiguous experimental results, it is generally concluded that this is due to problems with the protocol or its execution, not an unforeseen variable. Ambiguous experimental results are seen as detrimental to progressing knowledge as the aim is to produce usable results that do not rely on further interpretation. A process of blind variation occurs when experimental protocols are varied to get unambiguous results that are useable for publications. In contrast, ambiguous results in particle physics are seen as an opportunity to understand the problem better. Knorr-Cetina says this strategy is employed when problems arise, suggesting that this is maybe not the first preferred epistemic strategy, but rather one relied on in certain circumstances.

Beyond the differences in the actual process of knowledge creation Knorr Cetina observes that knowledge creation for each discipline takes place within different organizational structures, which dictates the division of labor, personal responsibilities, and professional recognition. For instance, the prevailing structure of molecular biology is small lab-based research groups, headed by a supervisor and composed of graduate students and post-docs. She also observes that various institutional pressures including career progression and funding could influence research direction. For instance, she found that post-docs were reported as being less keen to work on high-risk projects with potentially high payoffs as they needed a higher degree of certainty that their projects would deliver good results as this was a requirement to obtain a faculty position. The different disciplines both also had set norms through which the knowledge was legitimized, such as presenting at certain conferences or publishing in certain journals.

Knorr Cetina notes a social dimension in the process of knowledge creation, with even casual conversations between lab members demonstrating what they saw as valid knowledge and what made a successful scientist in their disciplines. This could be a way in which many of the casual disciplinary norms or assumptions are enforced. Knorr Cetina therefore finds that differences between disciplines have many different layers, including epistemic strategies, organizational structures, and social dimensions. She appears to agree with Flek in visualizing knowledge creation as a collective endeavor that is partly the product of the social climate it takes place in.

2.6 Peter Haas

Peter Haas is a political scientist. He achieved his PhD at MIT and is currently Professor of Political Science at the University of Massachusetts Amherst. His research interests include the role of science in global politics and epistemology. Haas is credited for developing the conceptual framework of epistemic communities, which he defines as a ‘network of professionals with recognized expertise and competence in a particular domain and an authoritative claim to policy-relevant knowledge within that domain or issue-area.' Haas proposes that epistemic communities emerge due to a top-down mandate for certain types of knowledge needed for policy.

Haas describes epistemic communities as having both knowledge and policy dimensions. According to Haas, what unites epistemic communities are their shared normative and principled beliefs, a shared notion of knowledge validity and a value-based rationale for community action and a common policy enterprise. Haas notes that people's concept of reality is mediated by their prior assumptions, expectations, and experiences, which can influence what knowledge they consider valid. A shared notion of knowledge validity is common among members of epistemic communities. Haas proposes that another uniting factor of epistemic communities is the way in which they use and advocate for the knowledge they produce. Within the community, these perspectives are reinforced through the members' institutional ties, informal networks and collective practices. This explains why Haas thinks that an epistemic community can span multiple disciplines, because policy opinions can more easily be shared by people across different disciplines, whilst for Knorr Cetina what defines epistemic cultures is the process of knowledge creation, including methodology and research process which are closely tied to the disciplinary practices.

It is important to note that Haas studies epistemic communities in the context of international policy making, whilst Knorr Cetina studied her epistemic cultures at the researcher level. As they view epistemic differences though different frames, it is understandable that they come to different conclusions on what defines them. Knorr-Cetina, Haas and Flek all consider how the social factors can influence the process of knowledge creation and how different people come to have different interpretations of what constitutes valid knowledge. Knorr-Cetina and Haas both make the argument that the process of knowledge creation cannot be understood unless the social and institutional factors are also considered. Haas for instance highlights the roles that different policy questions have in forming epistemic communities, whilst Knorr-Cetina notes how disciplinary norms and institutional and funding pressures can impact research direction. Although neither Knorr-Cetina and Haas explicitly refer to systems thinking, these examples could be interpreted as considering the process of knowledge creation to take place within certain systems, such as policy systems, institutional systems, or the system of western science. It is therefore important to consider whether systems thinking can provide insight into why different ways of knowing exist and consider if systems thinking is a viable way of understanding and characterizing potential knowledge systems.

2.7 Donnella Meadows

Donnella Meadows (1941-2001) was an environmental scientist, researcher and educator most well known for being lead author of the 1972 *The Limits to Growth* report and her 1999 posthumously published book *Thinking in Systems*, which was written to explain many of the concepts behind systems thinking so that readers could apply them to their own fields.

Systems thinking can both be a way of knowing, and a way of thinking about other ways of knowing, giving it a somewhat meta nature. Systems thinking has varied definitions but is mostly understood to be a way of viewing systems in a holistic manner, not in a reductionist (breaking it down into its consistent parts) or positivist sense (purely derived from empirical evidence.) Systems thinking appreciates that a part of a system may act differently when it is part of a larger system. Systems thinking is not a way of thinking exclusive to epistemology and can be applied in many contexts.

Meadows defines a system as ‘a set of things – interconnected in such a way that they produce their own pattern of behavior over time.’ Systems can have multiple individual components that can link together in different ways; including stocks, flows and feedback loops. The system can change or look different under certain pressures, including changing information flows, different rules of operation or the system taking place within an established paradigm. When these components combine, they can give rise to complex higher-level properties and emergent behaviors that cannot necessarily be determined by studying the individual components.

Systems thinking is a way of conceptualizing any set of things as a system. Although systems thinking has been proposed as the opposite to reductionism, Meadows clarifies that they can be complementary approaches. As systems thinking is a form of conceptualization, the components and boundaries of the system are not static but can be different depending on the frame used to view the system. This is one of the ways in which systems thinking overlaps with other ways of knowing.

A relevant systems thinking concept for epistemology are information flows through systems. Kuhn and Hacking appeared to describe paradigms and styles of knowing as consisting of somewhat static system components. In contrast, Meadows describes information and knowledge as something dynamic that can flow through a system and can be stored, transferred and transformed depending on the frame, perspective or component. Meadows describes how systems can have emergent properties, including abstract concepts such as knowledge. She also describes how systems can be governed by certain rules that in turn can impact the emergent properties of the system. Although he does not refer directly to systems thinking, Haas does appear to give examples of both of the above, most notably by explaining how epistemic communities advocating for certain knowledge and policy action emerge partially to aid policy makers in addressing specific uncertainties. Therefore, in this case the policy mandates act as system rules that in turn dictate the emergent property of the system, the different epistemic communities.

Chapter 3: Methodology and methods

This chapter provides a theoretical background to the methods used in this thesis.

3.1 Bibliometrics

Bibliometrics is a statistical method commonly applied within information sciences to identify the links between published works (Ball, 2018). Bibliometric analysis can take place at many different levels, including individual authors, papers or journals. The most common way of establishing such links is through citations or key words.

When the unit of analysis is the paper, papers can be sorted or ranked by metrics such as total number of citations, mean number of citations per year or citation half-life. They can also be sorted or ranked into different categories including country of origin or discipline, or key words.

Citations are, at their most basic level, an acknowledgment of the work of another expert. Bibliometrics research assumes that the number of times a paper is cited to some degree corresponds to the quality of the work. A paper cited more frequently is acknowledged as more influential in their field.

However, there are considerable limitations with using citation analysis as an analytical tool. Garfield (1979) describes how citation analysis can exclude potentially relevant factors including the varying contributions of individual authors on co-authored papers, author self-citations and citations to refute the conclusions of another paper. Hicks (1987) describes the large variations in citation behaviors that can occur at the disciplinary level. Citation does also not necessarily equate to the quality of the work, both because of the use of citations to refute (Onodera & Yoshikane, 2015) and that quality is a complex concept that cannot be reduced to a single indicator (Ricker, 2017).

In spite of these weaknesses, bibliometrics can be used for a number of functions, including identifying influential works and measuring publication output with respect to different variables (e.g. country, journals). A specific technique used in bibliometrics is clustering, which can generate more sophisticated insights about a research field.

3.1.1 Bibliometric Clustering

Cluster analysis looks at the degree of closeness between two objects with respect to the desired metric(s) (Everitt, 2011). In bibliometrics, these objects (or unit of analysis) can, for example, include papers, authors or journals. The metrics can include citations or key words, and there are various clustering methods. Cluster algorithms display the objects in space with the closest objects with reference to the desired metrics being spatially closer. The resulting visualization can be used to create a map of the objects, with different objects potentially being in distinct clusters.

There are various different ways to cluster objects, and choices need to be actively made by the researcher depending on the data set. For example, clustering can be done with respect to one variable or multiple variables weighted different ways. Objects can be sorted into clusters by average linking or by single links. Everitt (2011) argues that pre-choosing variables can enforce a cluster structure on the data. Good cluster analysis allows the 'natural 'clusters to emerge from the data without forcing a cluster structure that may not be good representations of the data.

Irrespective of method, variables used for bibliometric clustering include author names, affiliations, key-words and journals. A common approach is to do a co-citation analysis, which looks at how often papers are cited together. If, for example, two papers are cited together in multiple papers then they will have a high co-citation index and are more likely be in the same cluster as compared to papers that have a low co-citation index. Clusters will therefore consist of papers that are commonly cited together.

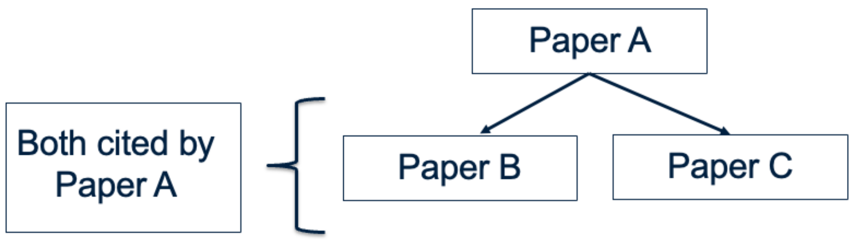

Figure 1 Illustration of co-citation analysis. As Paper B and Paper C are both cited by paper A, they will share at least one co-citation link

Another common approach is to analyze key word co-occurrence, which clusters papers by the number of key words they have in common. Clusters will be made up of papers that share similar key words. Therefore, these papers are likely to cover similar topics, themes or be in the same discipline. Such cluster maps can help identify research focuses or priorities in the field, including how this has changed over time.

Both of these approaches can offer more insights to a research landscape as compared to sorting or categorizing alone. Cluster analysis can be used to analyze which clusters are more prominent than others, which clusters are closer to each other than others, which papers link clusters or are prominent in their clusters. Hicks (1987) concludes that there are some limitations with co-citation clustering, as experts they interviewed in their study said that co-citation clusters were not reflective of the field as they perceived it. Hicks also note that differences in expert definitions made it difficult to apply co-citation analysis to policy applications.

3.1.2 Bibliometrics in the endocrine disruptor field

Bibliometrics has been used to study the ED field. For example, Miró et al., (2009) used bibliometric analysis to study publication output by toxicologists and in the field of toxicology. They explain the importance of publishing output in this field as it is commonly used to allocate funding.

Yuan et al. (2020) conducted a bibliometric analysis of Bisphenol A (BPA) and thyroid hormones. The aim of their study was to identify research hotspots and the future focus of the field. They sourced papers in this area through a key word database search (that included the term endocrine disruptor) and clustered them through key word co-occurrence as defined by the authors and journals. They found that publications in this area had increased significantly since 2000, indicating that it is a growing and active research field. The research hotspots they found included mechanistic research in animal models and epidemiological studies. This indicates that there might be distinct research clusters present in endocrine disruptor literature, and that they could some degree be differentiated by the differences in methods. However, it is important to note that the key words in the study were defined by the authors and journals. As discussed in the previous section, the norms in certain fields or disciplines could influence the terms that authors and journals use in their key word selection; meaning that authors could mean different things by the same term or use different terms for similar concepts. This demonstrates the importance of using a wide range of key words when using bibliometrics to source data.

Nayak et al. (2021) also conducted a topic-focused bibliometric search but on the adverse health effects of cosmetics. They used ‘cosmetics’ and ‘adverse effects’ as their key words for paper selection from the Scopus database. They created a bibliometric network visualization of the key words found in the study, which showed which key words generally appeared together, and which key words appeared most frequently. Bibliometric network visualization of key words can show which key words tend to co-occur and therefore what research areas tend to overlap. However, bibliometric key word analysis on its own does not give additional information on how the key words are used or contextualized in the field and without participatory feedback from study authors, we are not able to understand their interpretation. The key words used in the search terms were prominently featured, as expected, which emphasizes the importance of choosing key words that are used by researchers as not to exclude relevant papers. For example, the authors sorted the papers by research area to identify the field share of the respective research areas, but they equated research area with discipline (e.g. medicine, environmental science) and assumed that a paper could only belong to one discipline. Without checking with researchers, it is unclear whether this is a correct assumption, particularly in a field as interdisciplinary as toxicology.

(Buter et al., 2006) explored creating a bibliometric map of the endocrine disruptor field, and then combined the map with a qualitative concept map created by an expert in the field. An ED expert was asked to draw a map of how they perceived the ED research landscape, using their own language and specifying the links between different research areas. This was then compared to the bibliometric map of the ED literature. The expert map was then mapped onto the bibliometric map to see how the two interrelated. The authors note that, in fields such as EDs, different experts may give different meanings to various terms and they may use different vocabularies, and that these could differ from the key words found in published works.

3.1.3 Bibliometrics and epistemology

As described in Chapter 2, several terms have been used to describe different ways of knowing that may be held by groups of people, such as thought collectives (Flek), paradigms (Kuhn), epistemic cultures (Knorr Cetina), styles of reasoning (Hacking), methodological strategies, (Lacey) and epistemic communities (Haas). In the overview of bibliometrics and epistemology below, I will be using the term used by the author I discuss, as each of these terms have their nuances and reflect how the author perceives and conveys the phenomenon.

De Bruijn and Gerrits (2018) set out to identify epistemic communities in urban self-organization through coding and bibliometric analysis. They undertook a database search and selected papers that had been cited more than 10 times for consideration. Their coded categories were as follows: 1) type of study, (2) conceptualization of self-organization, (3) the issue(s) or topic(s) the concept of self-organization is applied to, (4) the method(s) deployed to research self-organization, (5) and the results of the process of self-organization. Choosing coding categories prior to analysis could have resulted in only discovering characteristics of the epistemic communities that were compatible with their prior assumptions. The authors identified two epistemic communities that differed in their conceptualizations of self-organization: with one community using it for modelling approaches and the other using it as a surrogate for self-governance. They then used bibliometrics to corroborate their findings. They found that co-citation coupling led to clusters that correlated with the epistemic communities observed.

3.1.4 Application of Bibliometric Clustering in the present study

To source papers researching EDs and human health impacts, a search was undertaken in the Web of Science database. The Web of Science database provides citation data for literature in many different disciplines. The search terms were designed to include all papers published between 1990 and September 2021 researching EDs and health impacts. Terms such as hormonal disrupt* were used to include papers where authors did not use the term 'endocrine disruptor'. This is because endocrine disruptor has been a politicized term since it was introduced by Theo Colburn in 1991, so whether an author chooses to use it or not could be indicative of an epistemic perspective. The full list of search terms used in the search are specified below:

> *((((((((((((TS=(endocrine disrupt* AND human health)) OR TS=(hormone disrupt* AND human health* AND pollut*)) OR ALL=(hormone disrupt* AND human health AND "chemical of emerging concern")) OR ALL=(hormone disrupt* AND human health AND "chemicals of emerging concern")) OR ALL=(hormonal disrupt* AND human health* AND pollut*)) OR ALL=(hormonal disrupt* AND human health AND "chemicals of emerging concern")) OR ALL=(hormonal disrupt* AND human health AND "chemical of emerging concern")) OR ALL=(hormone disrupt* AND human health AND contaminant)) OR ALL=(hormonal disrupt* AND human health AND contaminant)) OR ALL=(hormone disrupt* AND human health AND "substance of emerging concern")) OR ALL=(hormonal disrupt* AND human health AND "substance of emerging concern")) OR ALL=(hormone disrupt* AND human health AND "substances of emerging concern")) OR ALL=(hormonal disrupt* AND human health AND "substances of emerging concern")*

The search rendered 4,104 scientific papers. Full references from each paper were extracted.

The papers were clustered in VoS viewer. Citations totaled 41,066. Papers were clustered through full counting with the cluster capped at 5000 papers, sorted by the greatest total link strength.

The co-citation analysis resulted in six distinct clusters. A paper cannot belong to multiple clusters although certain papers will have different degrees of membership to multiple clusters. The occurrence of distinct clusters both shows that there are distinct groups of papers that cite each other which could indicate the distinct research communities. The co-citation clustering is assumed to provide an initial grouping of papers who acknowledge each other's work through citations.

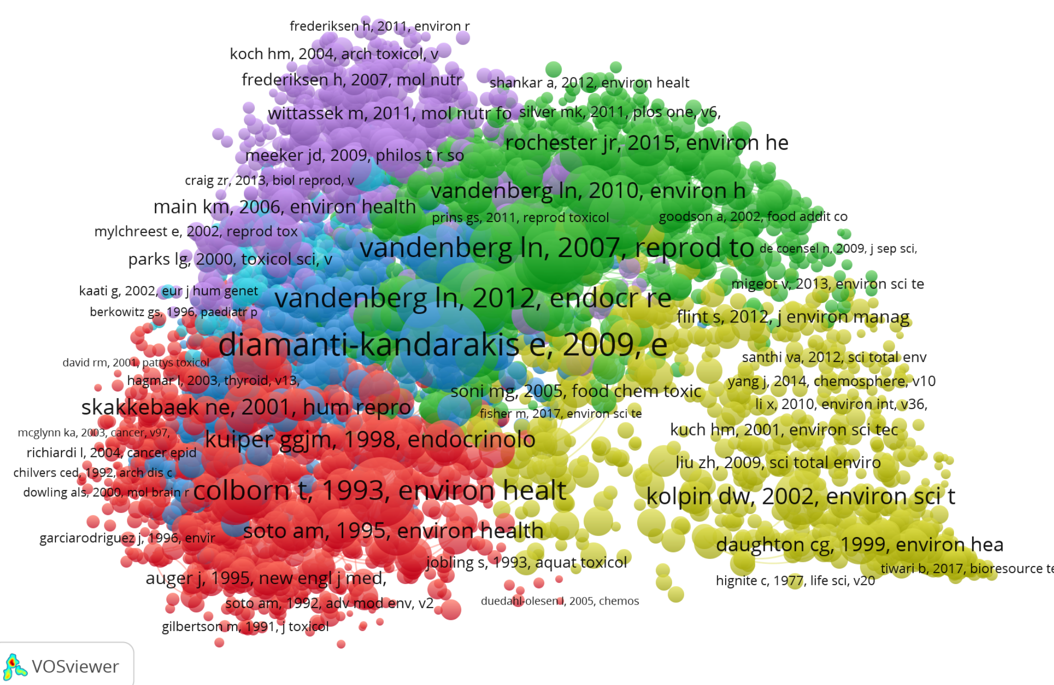

Figure 2 Bibliometric co-citation clusters

The top 10 most highly cited papers in each cluster (60 papers total) were selected for a preliminary read through. One paper was discarded from analysis as it described a PCR method which is irrelevant to this analysis and was not replaced. The 59 analyzed papers included various types of review papers and research papers using a variety of methods (epidemiological studies, assay tests, animal tests). From the initial read through, it was found that some papers in this initial paper selection tended not to just summarize research developments but set out to gather data together to support a hypothesis, suggest a hypothesis or advocate for a certain approach.

A preliminary read through the 59 papers guided the selection of papers which were assessed as presenting or making a clear argument. This was determined based on the identification of strong statements about hypotheses, conclusions or evidence that can be argued against. In some cases, the strength of the vocabulary (in bold) is not always proportional to the strength of the evidence cited. For example:

> ***It is known*** *that humans whose diabetes is being treated with rosiglitazone (a drug that activates PPARγ) develop more adipocytes and gain weight [22]. Therefore it is reasonable to hypothesize that chemicals capable of activating PPARγ might have the same effect (Schug et al., 2011, p. 207)*

> *Epidemiological studies* ***may be the "gold standard"*** *for assessing risks to human health, but these are expensive (costing upwards of one-million US dollars per study) and slow (typically a minimum of five years from conception of a human health study to results). (Swan, 2008, p.183)*

As grounded theory is a comparative method the papers used for analysis needed to demonstrate some degree of differences in reasoning. 13 papers were selected for the analysis. Two papers that met the identifiable reasoning criterion were taken from each of five of the clusters, and three from the remaining cluster (dark blue in figure 2). This was because the initial read through and analysis of the two papers from this cluster indicated that the papers in this cluster were less explicit in their reasoning, so another paper was required to strengthen the analysis.

3.2 Grounded Theory

Grounded theory is a qualitative method. The approach was originally developed by Glasser and Strauss in 1967, who used it to study interactions with terminally ill patients in hospitals. The rationale for grounded theory is that instead of developing a hypothesis and testing it on the data, the hypothesis is developed in tandem with analysis of the data. Therefore, the aim is not to test pre-conceived hypotheses or theoretical frameworks but to develop theories based on the findings in the data. A basic premise of Grounded theory is that the findings should not be presented as final or certain, but as the best theory developed from the sampled data. The underlying assumption is that this approach reduces the impact of the researcher's preconceived ideas or theories on the interpretation and analysis of the data, and thus the risk of seeing what one wants to see. It also allows for the study of areas for which there are limited literature and theories (Pidgeon, 1996).

Similar to many other qualitative methods, grounded theory does not just aim to describe the data, but to develop a conceptualization or deeper explanation for what the data shows. It is the difference between for example, describing that scientists say different specific things about a certain method, and explaining that there seems to be two distinct ways in which different groups of scientist conceptualize the same method.

Constructionist Grounded Theory (CGT) was developed as a stream of grounded theory by a student of Glaser and Strauss (Charmaz, 2000; Mills et al., 2006). CGT assumes that the researcher and the data cannot be separated; the researcher will to some degree always place their own assumptions and preconceptions on the data. This stream aims to acknowledge the researchers’ preconceptions but make them transparent whilst ensuring the findings are still grounded in the data (Ramalho et al., 2015).

How grounded theory is applied depends on the data available. For example, grounded theory can be applied to interviews as well as literature. A combination of data sources can contribute to a grounded theory. There are various specific methods developed to apply grounded theory, for example matrix based methods (Scott, 2004) and coding based methods (Charmaz, 2014). The method used in this study is based on (Charmaz, 2014) and will be described in the next section. The core characteristic of grounded theory can be summarized as representation, iteration, and validation.

Firstly, the data selected must be an appropriate source to study the chosen subject. If findings indicate that additional data is required to develop or clarify findings, then further data can be obtained at any stage of the investigation. The nature of grounded theory is iterative. Theories and findings change if the data shows that a finding is contradicted or more nuanced than previously found. The research process is not a linear progression from data to findings, but one that allows new findings to emerge at any stage of the research process.

Coding in grounded theory can “read beyond the text” (Charmaz, 2014): just because the researcher has an idea but cannot find concrete evidence for it in the text does not mean it has to be dismissed at the time. Therefore, inferences can be made from the available data without overreaching by reasoning beyond what is observable in the text, such as inferring author intention or perceptions.

As a last step, findings will usually undergo validation. Ways of achieving this can include multiple researchers independently coding and evaluating the findings collectively, or the researcher noting memos and observations throughout the research process and re-scrutinizing them later. Although the findings from a study cannot be certain, validation adds some rigor by seeing if the findings are truly grounded in the data. Validation can also help identify if saturation has been reached. This means that the codes provide a good enough representation of the data.

3.2.1 Grounded Theory and Epistemology

There have been qualitative attempts to explicitly characterize epistemic perspectives. Eyles et al. (2009) conducted an analysis using policy documents and expert interviews to identify the features of the heart health promotion epistemic community in Canada. The analysis aimed to identify common themes. Grounded theory findings should be specific, nuanced, and prevent prior assumptions from impacting the analysis. However, the grounded theory characterizations were very broad and reflected commonly held concerns, including the contextual factors of each provincial health system.

3.3 Application of Grounded Theory in the present study

Coding was conducted in Nvivo. This was done in line with constructionist grounded theory principles so was open and iterative. The first step was annotation coding, meaning that I set out to highlight quotes and sections of interest in these papers, while recording my interpretation of what they are portraying, while no attempts were at this stage made to categorize. These codes were used to support reading between the lines; noting what the authors appeared to imply through the text as well as explicitly state. I also used these codes to analyze the language used to express arguments; for instance, whether strong agreement or disagreement with a concept is expressed or present a balanced view. For example, Rochester (2013) asserts that “BPA is a known endocrine disruptor; it has been found to bind to estrogen receptors and have estrogenic effects in laboratory studies.” This was read as indicating that the writing that BPA is an endocrine disruptor is a factual statement and not something to be debated, as supported by the qualifier ‘known’ to endocrine disruptor. Jirtle & Skinner, (2007) assert that “Epidemiological evidence increasingly suggests that environmental exposures early in development have a role in susceptibility to disease in later life,” (p. 1) indicating that this paper has not reached a clear conclusion on the role of environmental exposures in human health.

As multiple papers were coded in succession, the codes were developed by comparing reasonings found in different papers. This was important as characterizing perspectives relies on identifying the differences in how different ideas were presented and expressed. Throughout this initial coding stage, notes were taken to record insights and reflections.

The next stage was axial, or category development coding. The aim of category development was to see which categories emerged from the data based on the themes identified through the annotation coding. Several initial categories were developed based on the different methods (e.g., epidemiological, animal testing) discussed in the papers, as the annotation coding suggested that the papers took different insights from these methods. Further coding and reflections led to argumentation categories which described how the papers articulated the use, interpretation, and application of certain types of knowledge, and how they compared methods to each other. As the previous individual categories did not allow for this comparison, I view the argumentation categories as more reflective of the data. These arguments were re-scrutinized and re-developed after discussion with my supervisor Gunilla Öberg.

3.4 Validation of identified arguments and testing the hypothesis that they form two distinct Ways of Reasoning

A follow-up study was designed to investigate if the identified arguments exist in other papers, to improve articulation of each argument by identifying nuances and re-scrutinizing the current arguments. It was also designed to systematically test the hypothesis that the arguments align to form the two hypothetically outlined WOR.

The most highly cited 10 papers were sampled from each of the six clusters that resulted from the bibliometric analysis (Figure 2). Of these 59 papers, 13 were used for the content analysis, and one was discarded as it was deemed to be irrelevant. The hypothesis was tested by analyzing the abstracts of the remaining 46 papers.

The definitions of the arguments identified and characterized during the analysis were used to prescriptively code the abstracts of the 46 sampled papers that were not used in the analysis. A co-occurrence analysis was then undertaken to identify the frequency with which the arguments occurred together in the same abstract.

3.4.1 A Note on the Author

I have a neurological condition known as synesthesia (Grossenbacher & Lovelace, 2001). I see letters, numbers and concepts as specific colors. Therefore, when I code I visually 'see' the different categories. I therefore kept coding until I achieved categories with distinct colors. As grounded theory relies on the researcher’s intuition, I want to acknowledge how I ‘see’ my intuition visually and use this in my study.

Chapter 4: Types of studies and arguments made in peer reviewed ED research

4.1 Types of studies in peer reviewed ED research

The analysis of the ED literature led to the identification of three major research approaches: Whole animal studies, mechanistic studies, and epidemiological studies. In the context of endocrine disruptors, whole animal studies commonly refer to any study that involves the direct exposure of a whole animal to a chemical combined with investigation into any potential effects or impacts on biological pathways. Mechanistic studies refer to any study that explores the biological pathways that link exposure and effect. Epidemiological studies investigate the association between exposure to (a) chemical(s) and adverse health outcomes in a population.

Mechanistic studies and whole animal studies can be undertaken separately and use distinct methods, but these approaches are often used in parallel and are often described in the present study as complementary approaches to establishing the mechanistic basis for effects. The analysis suggests that studies that involve animals can be divided in two categories based on their aim: the first seeks to test the effects chemical exposure has on whole animals, and the second aims to explain why. I have chosen to call the first whole animal studies and the second animal-mechanistic studies.

In whole animal studies a substance is administered to living animals, then, after a certain time following start of exposure, one or more variables are measured. For example, Hayes et al. (2010) exposed male frogs to atrazine leading to exposed males producing viable eggs.

Animal studies and mechanistic studies use animal models to understand the biological pathways resulting from chemical exposure and the effect at the level of a whole living organism. These studies are differentiated from whole animal studies because they attempt to establish and characterise the biological pathways underpinning health effects, either instead or in addition to purely characterising the effects from exposure. Various approaches can be taken in these studies. In some cases, observed effects in animals can be used to propose mechanistic explanations or an empirically investigated mechanism in an animal can be used to propose potential effects. For example, Somm et al. (2009) exposed pregnant rats orally to BPA from day 6 of gestation through to the end of lactation. It was found that perinatal BPA exposure predisposed rats to being overweight in a sex- and diet-dependent manner. These changes were associated with adipocyte hypertrophy and overexpression of lipogenic genes. This study was used to investigate the effect of exposure and develop an understanding of the cellular pathways linking exposure and effect, in other words a mechanistic rationale.

Mechanistic studies refer to any study that explores the biological pathways that link exposure and effect independent of a whole organism. These studies can be conducted using cells, subcellular structures (such as enzymes and DNA). These can include studies that investigate how chemicals interact with hormonal receptors and enzymes or genes that meditate hormonal activities and the impact of exposure on cell activity (e.g., cell proliferation, gene expression). For example, Raun Andersen et al. (2002) tested 24 pesticides for interaction with the estrogen receptor (ER) and the androgen receptor (AR) in breast cancer cell transactivation assays to test pesticides for their ability to bind to or block receptors. The pesticides were also tested for their ability to inhibit CYP19 aromatase activity and cell proliferation.

In the context of endocrine disruptor research, epidemiological studies are human or animal population studies that explore the links between exposure to endocrine disrupting compounds and health impacts on the population as a whole. Studies can do this for individual chemicals (Vandenberg et al., 2007), classes or groups of chemicals that are either structurally similar or have similar uses (Myers et al., 2016), endocrine disruptors as a class of chemicals (Dodson et al., 2012), or environmental chemicals which include endocrine disruptors (Taylor et al., 2018).

These studies commonly include measurements of exposure and health impacts. Epidemiological ED studies can have different design, structure, and timescales for measuring these variables. Some studies measure exposure and health impact at the same time, for example by measuring the concentration of chemicals in urine, blood, or in the environment that people currently live/work in (Liu et al., 2015). These are generally termed cross sectional studies and make comparisons at a single point in time. For example, Lee et al (2007) compared the serum concentration of persistent organic pollutant with the incidence of arthritis. Li et al (2011) compare semen quality between groups of men who have and do not have occupational BPA exposure.

Other studies measure exposure, then measure health impacts over a period of time. These are usually termed longitudinal or cohort studies. For example, Rundle et al. (2012) measured the blood concentration of polycyclic aromatic hydrocarbon in a cohort of pregnant mothers, and then measured the incidence of obesity in their children at 5 (n=442) and 7 (n=341) years old.

When it comes to health impacts, some studies look into one specific impact, while others consider multiple impacts. The impacts studies can include specific diseases (*e.g.,* breast cancer) (Liu et al., 2015; Sprague et al., 2013; Yang et al., 2009) or complex conditions with multiple influences (*e.g.*, obesity) (Carwile & Michels, 2011; Hatch et al., 2010; Trasande et al., 2009). Measurements can include incidence of a certain disease, biological measurements (*e.g.*, BMI, anogenital distance) or observed pathologies (*e.g.*, reproductive function, ADHD).

The analysis led to the identification of 13 arguments, summarized in Table 1. Of these seven were purely about mechanistic studies, three purely about epidemiological studies and two about both mechanistic and epidemiological studies. This refers to the arguments reasoning about the validity and function of certain methods.

Table 1. Overview of the 13 arguments identified in the present study in the order they are discussed, the acronym used for these and whether they are related to mechanistic or epidemiological studies or both

#	**Argument**	**Acronym**	**About**
1	Inherent Toxicity	ITA	Mechanistic studies
2	Mechanistic Causality	MCA	Mechanistic studies
3	Mechanistic Extrapolation	MEA	Mechanistic studies
4	Analogous Effect Argument	AEA	Mechanistic studies
5	Individual Chemical	ICA	Mechanistic studies
6	Contextuality	CA	Mechanistic studies and Epidemiogical Studies
7	Reductive Mechanistic	RMA	Mechanistic studies
8	Collective Harm	CHA	Mechanistic studies and Epidemiogical Studies
9	Difficulty in Extrapolation	DEA	Mechanistic studies
10	Epidemiological Characterization	EChA	Epidemiological studies
11	Epidemiological Harm	EHA	Epidemiological studies

12	Dual Evidence Causality	DECA	Mechanistic studies and Epidemiogical Studies
13	Lack of Association Argument	LAA	Epidemiological studies

4.2 Arguments relating to mechanistic studies

The analysis led to the identification of arguments that were broadly supportive or critical of mechanistic studies to understand the human health effects of EDs. Each argument is distinct, but they interlink as certain arguments rely on others for their reasoning. Arguments supportive of mechanistic studies hold that these studies can provide valuable insights critical to understanding the hazards associated with EDs. Arguments critical of mechanistic studies point out that mechanistic studies have limitations as it pertains to evaluating chemical risk. These arguments also assert that these limitations are not sufficiently acknowledged when results from mechanistic studies are used to support formalized chemicals risk assessment. The term risk assessment I am referring to here, is the codified evaluation procedure used in regulatory frameworks such as under the Canadian Chemicals Management Plan to assess the potential adverse health effects a chemical can cause and the likelihood of these effects occurring.

4.2.1 The Inherent Toxicity Argument (ITA): by studying chemicals individually, you can remove the contextual factors and pin down inherent toxicity

The present analysis led to the identification of an argument hereafter called ‘The Inherent Toxicity Argument (ITA)’ which has as its core two claims: a chemical with endocrine disrupting properties is inherently hazardous and the strength with which a chemical produces a degree of endocrine disrupting effect varies among chemicals.

From this argument, it follows that inherent properties are best evaluated at the individual chemical level. This is because, contextual factors will disrupt the ability to 'see' the inherent properties by introducing multiple variables for consideration. Contextual factors are consequently less relevant, or even undesirable to consider. This line of argument implies that certain methods are better suited to assess a chemical's inherent activity than others. Whole animal and mechanistic studies do, by their design, allow for the study of individual chemicals and isolation from other factors that may make the chemical's inherent toxicity more difficult to discern.

Controlled exposure studies are central to the ITA, as these types of studies are used to produce evidence that can directly link an effect to a chemical. The basis for this argument is that knowing that the animal has only been exposed to this chemical and that other factors are either eliminated or controlled, makes it possible to make the assertion that the investigated chemical is responsible for an observed health impact. For example, in the excerpt below, it is asserted that controlled animal and human exposure studies give a higher confidence than epidemiological studies in identifying cause-effect relationships due to the controlled exposure conditions. As compared to epidemiological studies, controlled exposure experiments render more certainty that the exposure is directly responsible for the health effect. The core claim here is that the design of epidemiological studies does not allow for the measurement of exposure variables and other potential causal factors, making it more difficult to isolate a pathway from exposure to effect that would allow the assertion of a direct causal link; or an assertion that the effect is due to the chemical's inherent toxicity, and not to other factors.

> *…because many epidemiological study-types (i.e. cohort, case control, cross-sectional studies) by their nature do not include controlled exposures, and many do not include measured exposures prior to outcome measurements, these studies are inherently less strong in terms of confidence, compared to animal studies and controlled human exposure studies. (Rochester, 2013, p. 139)*

Mechanistic studies are seen as the 'gold standard' method under this argument. This is firstly because mechanistic studies are thought to provide a causal explanation, because mechanistic activity is independent of other factors whereas other methods are not. Mechanistic activity is viewed as mostly structurally determined, even though environmental factors can moderate it. For example, Swan (2008) asserts that understanding the mechanism is a key component of causality.

> *Moreover, epidemiology is limited in its ability to draw causal inferences. Therefore, it must be informed by, and closely linked to, animal studies in which mechanism and causality can be demonstrated. (Swan, 2008, p. 183)*

Through this argument, it is important to know which mechanisms lead to health effects as this provides a proof that the chemical's inherent toxicity is responsible for any effects, and this is not influenced by any contextual factors.

4.2.2 The Mechanistic Causality Argument (MCA): If the mechanisms that cause health effects are known, it is possible to make inferences about health effects

The Mechanistic Causality Argument (MCA) builds on ITA in that it asserts that knowing the mechanisms through which endocrine disrupting chemicals act in the body can be used to make inferences about the health impact of chemicals. The argument consists of two interlinked parts.

The first part asserts that mechanisms are seen as the key to identifying causal links between exposure and health effects because they provide a rationale for why effects happen, not just that they do. According to this part of the argument, it is not enough to state that effects happen to conclude that certain chemicals are likely to cause harm because the rationale (i.e. causal links) behind the effects must be understood. In keeping with this argument, mechanistic studies are an indispensable component when making inferences about the potential harm that a chemical may cause. This line of reasoning is for example seen in Swan (2008) who asserts that understanding the mechanism is a key component of causality.

> *Moreover, epidemiology is limited in its ability to draw causal inferences. Therefore, it must be informed by, and closely linked to, animal studies in which mechanism and causality can be demonstrated. (Swan, 2008, p. 183)*

Similarly, Rochester (2013) asserts that it is important to demonstrate the "biological plausibility" (p. 143) of health effects through mechanistic studies and the paper considers it of great importance to establish the molecular mechanisms mediating the effects. Grün & Blumberg (2006) contend that a "plausible molecular mechanism" (p. 50) is desired to explain the observed effect and upholds that mechanistic evidence is needed to "support this model" (p. 50).

The second part of the MCA builds on the similarities between natural and synthetic hormones. This part of the argument asserts that the roles and pathways of the body's natural hormones are quite well understood. Therefore, if a synthetic ED has a structural similarity to a 'natural' hormone then it is reasonable to assume that synthetic EDs will follow the same pathways as natural hormones, and thus interfere with their regulation. For example, Grün & Blumberg (2006) explain that their hypothesis for the effects of chemical exposure parallels the known pathways of natural hormones that are structurally similar. Similarly, as seen in the excerpt below, Schug et al (2011) alludes to the inherent properties of EDs being similar to natural hormones leads them to have similar modes of action within the body.

> *EDCs are capable of mimicking natural hormones and maintain similar modes of action, transport, and storage within tissues. The properties of these chemicals, while unintended, make them particularly well suited for activating or antagonizing nuclear hormone receptors. (Schug et al., 2011, p. 205)*

Combining the two parts of the argument allows for the assertion that it is possible to extrapolate from mechanistic studies, commonly carried out on animals, to health effects on humans. In other words, the causal relationship observed between a specific chemical and a specific effect, can be used as a proxy for the health effects in humans, meaning mechanistic knowledge can be used to propose human health effects. The first part of the argument asserts that mechanisms underly health impacts, whilst the second part says that structural properties of the endocrine disruptor can lead to predictions about which mechanisms the ED can interfere in. For example, Hamers et al. (2006) concludes that EDs with certain mechanistic profiles are more likely to have health and environmental implications than EDs with other mechanistic profiles, stating that "implications for human and environmental health are expected for BFRs grouped in clusters I (TTR-binding and E2SULT inhibition)." (p. 171)

Kuiper et al. (1998) proposes that certain EDs "may trigger many of the biological responses that are evoked by the physiological estrogens." (p. 4252) Therefore, it is assumed that these EDs will act as estrogens, which can lead to inferences about the type of health impacts that will result from exposure. Similarly, Schug et al. (2011) asserts that chemicals that act through similar mechanisms may result in similar effects.

> *It is known that humans whose diabetes is being treated with rosiglitazone (a drug that activates PPARγ) develop more adipocytes and gain weight [22]. Therefore it is reasonable to hypothesize that chemicals capable of activating PPARγ might have the same effect (Schug et al., 2011, p. 207)*

In conclusion, MCA allows inferences to be made about the health effects of chemicals based on their mechanistic behavior. The justification is based on the reasoning that mechanistic evidence is the gold standard for causal claims between exposures and effects, and the structural and mechanistic resemblances between synthetic and natural endocrine disruptors.

4.2.3 Mechanistic Extrapolation Argument (MEA): If the mechanisms underlying a health effect in one organism is known, these findings can be extrapolated to other organisms

This argument is similar to the MCA and could be considered to be a component of the MCA, but due to the importance given to extrapolation in the papers it is considered here as a separate argument.

Similar to the MCA, the Mechanistic Extrapolation Argument (MEA) argument builds on ITA in that it assumes that information about mechanisms can be used to causally explain health effects. MCA and MEA are complementary and are at times used in parallel and at times used separately. Both hold that mechanistic knowledge allows for extrapolation. While MCA holds that one can extrapolate from one chemical to another, MEA asserts that it is possible to extrapolate from one organism to another, even if the contextual factors are different. This allows for the conclusion that a chemical that acts a certain way biologically in an animal will act similarly in humans. This relies on the core of ITA, that the mechanism is inherently determined and is independent of contextual factors, although the effect can be moderated by such factors.

For example, as illustrated in the excerpt below, Kuiper et al. (1998) argues that insect cells are a suitable model to assess ED mechanistic effects in mammals as the process in insect cells “closely parallels” (p. 4255) the process in mammalian cells. The mechanistic findings through one model (insect models) can therefore be used to extrapolate to another model (human models) and remain relevant.

> *Furthermore, it has been demonstrated that posttranslational processing of proteins produced in Sf9 insect cells closely parallels these events in mammalian cells (53). It was therefore decided to use human ER-alpha and ER-beta protein expressed in insect cells for the ligand binding experiments. (Kuiper, 1998, year, p. 4255)*

In a similar vein, Casals-Casas & Desvergne (2011) holds that animals “provide a reasonable approximation of human metabolism for most compounds” (p. 153). According to this line of argument, as the biological pathways in animals are similar to humans, it can be inferred that a chemical will act similarly in humans, in other words, extrapolation from one type of organism to another is appropriate.

4.2.4 The Analogous Effect Argument (AEA): findings can be extrapolated between species

AEA, as it is conceived here, takes the line of reasoning that forms the basis for MEA a step further by assuming that if mechanisms are similar in different animals, and mechanisms are the underlying cause of health effects, then it is possible to conclude that similar health effects observed in different organisms can have the same underlying cause, which means findings in one species can be relevant to another.

The idea of analogous effects is of great importance to this argument. The term analogous effects refers to the same adverse effect from the same chemical exposure being expressed differently between organisms. This argument assumes that analogous effects have the same underlying cause and differences in presentation are mostly due to differences in species. For example, Swan references the 'human analogue' of phthalate syndrome (p. 177)

If indeed the underlying mechanism is found to be the same, a chemical could pose a similar hazard to animals as they pose to humans. Because effects can be analogous, if an effect resulting from chemical exposure can be found in one organism inferences can be made on how the same chemical can affect different organisms, including humans.

For example, Colborn et al. (1993) found that children and the rats exposed to the same chemical(s) display similar effects. They use this observation to strengthen the claim that these chemicals are causing health effects. Kuiper et al. (1998) uses observed effects in wildlife to propose similar effects to humans exposed to the same chemical.

> *The childrens' intractable behavior appears to be analogous to the behavior of the rats fed Lake Ontario fish. (Colborn et al., 1993, p. 380)*
>
> *Threats to the reproductive capabilities of birds, fish, and reptiles have become evident and similar effects in humans have been proposed (Kuiper et al., 1998, p. 4253)*

Colborn et al. (1993) stresses that whilst one species provides the model for maternal transfer, experiments in other species were able to confirm this finding under controlled conditions. They use this line of argument to support the idea that if the chemical exposure and mechanistic model is the same, findings can be extrapolated between species.

> *wildlife species have provided the model for maternal transfer of environmental endocrine-disrupting chemicals with their resulting suite of*

> *effects in offspring; experiments with laboratory animals have confirmed the findings. (Colborn et al., 1993, p. 381)*

4.2.5 The Individual Chemical Argument (ICA): chemicals are best evaluated at the individual level

The core of ICA is that it is better to study chemicals at an individual level because mechanistic activity is structurally determined, potential hazards can be attributed to individual chemicals, and some chemicals are more hazardous than others. This argument posits that greater hazards translate to greater potential for a chemical to cause harm.

The first part of this argument asserts that individual chemicals have individual hazards. These hazards are viewed as being inherent to the chemical and mechanistic activity can be used to infer potential health impacts.

The first stream of this argument implies that direct causal links between the potential harm posed by a chemical can only be made in relation to individual chemicals. For example, Kabir et al. (2015) argues that you can only identify causal links between individual chemicals and effects, as it is impossible to design studies that capture persistent and widespread exposure to mixtures.

> *However, although there are several examples of toxicities and contamination from PCBs and dioxins which may show a direct casual relationship between a chemical and manifestation of an endocrine or reproductive dysfunction in humans and wildlife, there is no representation of a common widespread persistent exposure to a broad mixture of contaminants. (Kabir, 2015, p. 245)*

Colborn et al. (1993) stresses the need to understand which chemicals are responsible for endocrine disrupting effects, implying that only some individual chemicals are responsible for ED effects, and these need to be identified.

> *This is due in part to the lack of knowledge concerning which of the many chemicals present in the environment are responsible for endocrine-disrupting effects. (Colborn et al., 1993, p.381)*

The next step in ICA builds on ITA, *i.e.*, that a chemical´s ability to cause an adverse effect is inherent to the chemical and to some degree structurally determined. The next part of this argument asserts that some chemicals are more toxic than others. This follows on from the first stream because assuming that some chemicals cause certain effects and others don't, or that some chemicals have structures that are hormonally more active than others, which leads to the conclusion that some chemicals will have more severe effect than others, which is inferred to translate into higher potential to cause harm to humans and other organisms.

For example, Hamers et al. (2006) asserts that chemicals with certain mechanistic profiles are expected to have human health implications. This implies that certain chemicals have higher potential to cause harm than others based on their mechanistic profiles.

> *implications for human and environmental health are expected for BFRs grouped in clusters I (TTR-binding and E2SULT inhibition) (Hamers, 2006, p. 171)*

ICA, in combination with the high number of chemicals in circulation and limited resources to address each one of them, is used to argue for methods that allow for rapid testing and assessment of large numbers of individual chemicals, such as high throughput assay testing.

For example, Kabir contends that, among the spectrum of ED chemicals, only a certain number can be taken for further analysis. Mechanistic studies can be used to select chemicals for further analysis that they deem likely to cause harm to humans and other organisms.

> *Furthermore, the analyst has to make a decision about how to best allocate analytical resources for this task. While the "spectrum of the EDC chemicals" (not isomers or congeners) is very large, only a subset can be extracted and separated by chromatography, and even less numbers identified. The use of effects directed analysis (EDA) and quantitative structure-activity (Kabir et al., 2015, p. 248)*

Hamers et al. (2006) argues along a similar line by alluding to the inherent hazards of EDs as a way of systematically identifying the most hazardous individual chemicals. They also argue for the policy relevance of this method, as this is how the EU wants to assess these chemicals.

> *The goal of the current study was to perform a systematic in vitro screening of the potential ED potencies of BFRs, as part of a hazard identification approach within the EU-funded research program FIRE (RIVM, 2003). (Hamers, 2006, p. 157)*

Casals-Casas raises that chemicals have to be dealt with individually as it is unclear how multiple chemicals can be treated together in regulation.

> *At a practical level it is also unclear how regulatory decisions will be amended to account for exposure to mixtures rather than to single compounds. (Casals-Casas, 2011, p.153)*

4.2.6 Contextuality Argument (CA): contextual factors are required to understand the risk of a chemical

CA asserts that contextual factors cannot be excluded as the intrinsic activity of the chemical cannot be separated from the contextual factors and limited inferences can be made based on intrinsic activity alone. This argument criticizes mechanistic studies because they think these studies assume that intrinsic activity that can be studied isolated from the system to make inferences about how it behaves within the system.

According to this argument it is important is understand the dynamics of exposure, meaning when, how and where people are exposed to chemicals. Contextual uncertainty is here seen as a more urgent concern than mechanistic uncertainty. This argument still emphasizes the usefulness of mechanistic studies in understanding effects or demonstrating endocrine disruptor activity but hold that such studies have limited value in inferring the potential harm as this requires assessing the consequences of exposure as it happens in the environment, it is consequently not possible to infer harm without calculating the risk, which requires considering contextual factors.

For example, Swan (2008) asserts that differences in contextual factors can lead to difficulties in interpreting results across species. They also stress the importance of chemical risk assessment being relevant to human exposure; if the contextual factors are not considered this decreases the relevance of the results.

> *Because of these important differences in route, dose level and multiplicity of agents, it is often unclear how to interpret findings across species. For these reasons, if chemical risk assessment is to be relevant to human exposure, future assessments should incorporate more environmentally relevant scenarios, including mixtures. (Swan, 2008, p. 183)*

Kabir (2015) holds that knowing the mechanisms on their own is not sufficient to draw conclusions about the risk. Knowledge of multiple contextual factors associated with exposure and metabolism is required to establish the role of an ED in a disease.

> *Though we know the mechanisms by which endocrine disrupting chemicals (EDCs) work in biological systems, we cannot say with certainty that EDC is involved for a particular disease. A number of challenges like unique*

exposure rate of different persons to a variety of known and unknown EDCs; individual differences in metabolism and body composition (which may lead to the variation of half-life, persistence and degradation of EDC in body fluids and tissues, etc.) as well as differences in genetic polymorphisms etc. may be required to be ascertained to determine their involvement. (Kabir, 2015, p.250)

Casals-Casas & Desvergne (2011) describes the complexity of exposure dynamics, implying that varying types of contextual knowledge related to metabolism and systems biology are required to assess the consequences of exposure.

New integrative approaches combining genomics, proteomics, metabolomics, systems biology, and computational modeling should help to understand the complexity of the cocktail effect and its consequences when exposure occurs at various life stages./…/ Although cellular models are flexible and allow large numbers of conditions and doses to be tested, they are limited and unable to factor in bioavailability of the compounds and their metabolites or the route of exposure. (Casals Casas, 2011, p. 154)

4.2.7 Reductive Mechanistic Argument (RMA): mechanistic studies oversimplify by excluding relevant contextual factors

The core of RMA is that mechanistic studies are reductive and do not tell you a whole lot about how hazards translate into risk. RMA says that mechanistic studies over-simplify by eliminating certain factors that need to be considered when making inferences about the potential harm that chemicals might cause to humans and other organisms.

Schug et al. (2011) explains that is not appropriate to make predictions of how safe chemicals are when the doses and endpoints used in testing studies are not reflective of actual endocrine disruptor effects and exposure.

In summary, making predictions about the safety of chemicals by testing at moderate or high doses is not appropriate when very low doses of

> *endocrine disruptors can alter biochemical and morphological endpoints in a manner that is not necessarily predicted by exposures at much higher doses. (Schug et al., 2011, p. 208)*

Casals-Casas & Desvergne (2011) hold that mechanistic studies oversimplify mechanistic pathways. This includes reducing mechanisms to receptor interactions or ignoring the action of multiple simultaneous pathways.

> *Reductionist molecular and cellular studies may not take into account the knowledge that several EDCs interfere with a variety of NRs, not all of which are expressed in the same tissue or at the same time. Additionally, EDCs can affect several organs in the body with different intensities, leading to a global response that may be opposite of that observed in cell cultures. (Casals Casas, 2011, p. 153)*

RMA stresses that when observed at different scales, the effects of EDs may look very different. This is in contrast to the MEA that asserts that if the underlying mechanisms are the same, then the effects can be inferred from these. RMA and MEA therefore differ on whether the mechanism can be viewed as being independent from the system, or not.

4.2.8 Collective Harm Argument (CHA): the potential harm posed by chemicals are collective

The core of this argument is that there is limited value in studying chemicals individually because the potential harms posed by EDs are collective as organisms are always exposed to several chemicals (meaning that EDs collectively contribute to health effects). Under this argument, the health effects caused by EDs are ascribed to EDs collectively, not individual chemicals. A central component of this argument is that EDs are present in the environment as mixtures. Kabir et al. (2015) for example, emphasizes that wildlife is not exposed to single contaminants, and Casals Casas & Desvergne (see excerpt above) refer to the ‘cocktail effect’ of chemical mixtures in the environment.

> *There are a number of evidences all of which have indicated that endocrine disruptors are also responsible for different wildlife crisis. However, wildlife is not exposed to single contaminants, but is exposed to a complex mixture of endocrine disruptors. (Kabir, 2015, p. 244)*

If EDs are present in the environment as mixtures, it follows that they have the potential to interact or act together synergistically to cause effects. Therefore, the causality of effects cannot be attributed to individual chemicals because multiple chemicals contribute to them collectively.

Schug (2011) refers to EDs collectively as a public health problem. They advocate that more characterization is needed on the specific health effects associated with EDs, rather than the relationship between one chemical and its effect that are more commonly evaluated in mechanistic studies and controlled animal exposure studies.

> *The notion that EDCs are significantly impacting human health is of great concern. More data is needed to expand the list of tissues affected by EDCs, and more effort is needed to identify and classify the diseases and dysfunctions they are causing in humans and animal models. Nevertheless, current data is sufficient to identify a public health problem that must be addressed. (Schug et al., 2011, p. 213)*

Kabir et el. support Schug because they argue that studying every chemical individually is both unachievable and counterproductive because it fails to include all chemicals that are potentially harmful.

Kabir et al. (2015) implies that it is unachievable to study every single chemical that has the potential to cause harm.

> *The analysis of target analytes can be viewed as a top-down approach that only scratches the surface of the number of chemicals that can be measured. (Kabir et al., 2015, p. 248)*

4.2.9 Difficulty in Extrapolation Argument (DEA): contextual factors impede extrapolation between species

The core of the Difficulty in Extrapolation Argument (DEA) is that differences in contextual factors make it difficult to extrapolate between species, which is one of the central tenets of mechanistic studies. While MEA asserts that as mechanisms are the underlying cause for any health effects, then, if there is alignment in the mechanism, findings can be extrapolated between different species. According to MEA, contextual factors are of less importance in considering how an effect comes to be. DEA disputes the MEA argument by asserting that these contextual factors have such a sway over how the effect the chemical causes that they can make extrapolation extremely difficult, if even possible.

The first justification of DEA is that controlled exposure studies render animal models that do not reflect actual human exposure in the environment. For example, Casals-Casas & Desvergne (2011) argues that the exposure and doses used in animal studies are not reflective of actual human exposure in the environment, specifically through chronic exposure. Similarly, Swan, (2008) expresses the difficulty of making links between human and animal studies due to the differences in testing.

> *Issues of dose and exposure time complicate experimental design. Mice may live for two years, and a three-month exposure can be considered chronic. Does this experimental setup accurately reflect the decades over which humans might be exposed? (Casals Casas, 2011, p. 153)*
>
> *While several links between rodent and human studies have been made in this area, as discussed above, these are limited by several discrepancies between human and animal testing. (Swan, 2008, p. 183)*

4.3 Arguments related to epidemiological studies

The following four arguments differ with regards to inferences, *i.e*., the type of conclusions they claim can be legitimately drawn from epidemiological studies (for example, causal links, suggested links); purpose, *i.e.,* the context where epidemiological studies can contribute (e.g. for risk evaluation, for effect characterization, for demonstrating causal links); and validity, *i.e.,* the factors that are held as necessary for epidemiological studies to count as 'good enough' for the specified purpose (e.g. replicated results, sample size). Some of the arguments presented here are not against epidemiological studies, but disagree with other arguments on each of these points.

4.3.1 Epidemiological Characterization Argument (EChA): epidemiological studies are a suitable means to study the effects of EDs

This argument holds that epidemiological studies provide the truest representation of how chemical exposure impact humans. This argument relies on four central reasons, as further outlined below.

The first reason relates to latent effects. EChA asserts that there can be a significant time delay between exposure and the points at which effects are observable. Consequently, as longitudinal studies by design allow for a time delay between exposure measurement and effects measurement, they are a suitable method to assess the impacts of latent effects. For example, Colbourn *et al*. (1993) argue that there are reasons to believe that the effects of ED exposure are not immediately observable, (*i.e*., effects are latent) and, with a suitable time delay, epidemiological cohort studies are therefore an adequate method to assess such effects.

According to this line of argument, cohort studies are sufficient to draw the conclusions that there is a link between exposure and effect.

> *The possible immunological and endocrinological consequences remain to be determined in these cohorts. A major concern is that some of these consequences may not become apparent until young adulthood or even middle age. (Colborn et al., 1993, p. 381)*

The second reason relates to exposure in epidemiological studies mirroring real life exposure. EChA asserts that human exposure to EDs is complex, with multiple chemicals, long periods of low-dose exposure and multiple exposure pathways. Epidemiological studies account for this complexity and take a holistic view of what happens when chemical exposure occurs. This frames ED health effects as inextricably linked to the environment. As environmental factors including exposure timing, route and dose play an important role in how health effects develop it is important to understand these factors.

For example, Casals-Casas & Desvergne (2011) and Schug et al. (2011) both stress the importance of understanding exposure factors to understand effects.

> *New integrative approaches combining genomics, proteomics, metabolomics, systems biology, and computational modeling should help to understand the complexity of the cocktail effect and its consequences when exposure occurs at various life stages. (Casals-Casas, 2011, p. 154)*
>
> *The duration and route of exposure may also have a big influence on how the chemical is metabolized and whether or not the chemical remains biologically active. Additionally, the "low dose" levels at which these chemicals function are lower than those typically used in standard toxicology testing. (Schug et al., 2011, p. 208)*

Schug et al. (2011) assert that it is the interactions between an individual and their environment that contribute to health effects. It is therefore important to understand these factors.

DOHaD (developmental origins of health and disease) describe the interactions between the developing organism and the environment that determine the propensity of that individual to develop disease later in life [1].' (Schug et al., 2011, p. 205)

The third reason relates to epidemiological studies enabling the measurement of endpoints that are not accounted for in other study types. This reason is based on the claim that certain effect endpoints for endocrine disruptors can be more subtle and not easily identifiable in other study types; both because they are less obvious without a large sample size, and that these endpoints result from long term exposure.

For example, Grün & Blumberg (2006) highlight that 'abnormal induction of male sex characteristics in female marine invertebrates represents one of the clearest examples of environmental endocrine disruption.' This is presumably because it is easy to observe, by being a dramatic change, and easily linked to the endocrine system. Other authors assert that the effects due to endocrine disruption can be less drastic, leading to them being undetected.

The possibility for continual but undetectable or unnoticed effects on aquatic organisms is particularly worrisome because effects could accumulate so slowly that major change goes undetected until the cumulative level of these effects finally cascades to irreversible change-change that would otherwise be attributed to natural adaptation or ecologic succession. (Daughton & Ternes, 1999, p. 907)

Colborn et al., (1993) points out that the manifestation of effects is not always obvious until organisms reach certain developmental stages.

Taking into consideration what is currently known about chemicals that disrupt the endocrine system, the effects 1) may be manifested in an entirely different way, and with permanent consequences, in the early embryo, fetus, and neonate from effects as a result of exposure only in adulthood; 2) can change the course of development and potential of offspring, with the outcome depending on the specific developmental period(s) of exposure; and 3) are often delayed and thus may not be fully or obviously expressed until the offspring reaches maturity or even middle

> *age, even though critical exposure occurred during early embryonic, fetal, or neonatal life. (Colborn et al., 1993, p. 378)*

Rochester (2013) stresses that epidemiological studies are capable at picking up subtle ED effects, particularly due to low doses.

> *When concentrations of endocrine active substances exceed the physiological amounts (as is often the case in 'traditional' toxicology studies) these sensitive endocrine effects can be attenuated [2]. Thus, incorporating these epidemiological studies, which examine the health effects in environmentally exposed humans, can help elucidate these low-dose, non-monotonic effects.' (Rochester, 2013, p. 151)*

Casals-Casas & Desvergne (2011) assert that even though cross-sectional studies cannot provide a causal link these studies still have value by demonstrating the effects of EDs on humans.

> *Most of these studies are cross-sectional and as such fail to provide a causal link but rather draw attention to the potential effects of PFCs on human physiology* (Casals-Casas & Desvergne, 2011, p. 142)

4.3.2 Dual Evidence Causality Argument (DECA): epidemiological studies are a suitable means to identify links between exposure and effect, but do not by themselves provide sufficient evidence

This argument holds that epidemiological studies are valuable but not in themselves sufficient to demonstrate causal links. It relies on two central reasons.

The first reason builds on the claim that mechanistic evidence is required for causality. Mechanistic evidence, in this context, is generally understood to be a demonstrated or speculative but plausible biological pathway through which an endocrine disruptor can act in the body that can lead to an effect manifesting. For example, an endocrine disruptor could be found to bind to estrogen receptors in an assay test. According to this line of reasoning, the type of evidence gathered from epidemiological studies is seen as weak evidence because it does not allow the development of a rationale as to *why* a health effect happens, not just that it does.

> *For example, cross-sectional studies lack any information on timing of exposure and include only prevalent cases. They can suggest a link between a disease and the factor of interest…cohort studies are stronger than cross-sectional and case-control studies when well executed, but they are more expensive…Due to the potential for many forms of bias in this study type, case-control studies provide relatively weak empirical evidence, even when properly executed. (Casals-Casas & Desvergne, 2011, p. 141)*

The second justification is that association on its own is not enough to assert causality. A causal link through this line of reasoning requires a known path from exposure to effect, including timing method and amount of exposure, and an isolation from other potential causal factors. As animal studies can control these variables to a better degree that epidemiological studies, those types of studies are capable of establishing causality.

For example, in their 2006 study of BFR compounds Grün & Blumberg (2006) write that epidemiological studies have "suggested a link" between exposure and effect. According to the reasoning in this paper, speculative hypotheses derived from epidemiological studies are useful as a as starting point for mechanistic studies. According to this line of reasoning, epidemiological studies are useful because they can be used to identify a potential link, but mechanistic studies are needed to establish causality. Accordingly, epidemiological studies cannot, by themselves, be used to draw the conclusion that an effect is linked to a certain exposure.

DEC holds animal studies and epidemiological studies complement each other. This assumes that exposure under controlled conditions and in different species and associated effects to some degree corresponds to human exposure and effects. On the assumption that this parallel can be made, epidemiological studies can be used as supporting proof towards the claim that a specific chemical causes the effect under actual conditions of human exposure.

For example, Rochester (2013) notes that it is difficult to make causal links based on epidemiological studies but they hold that the increasing burden of correlative evidence combined with evidence from animal studies is sufficient to make a tentative assertion of the potential harm. In this case, epidemiology and animal studies can be complementary in assessing chemical risk.

> *While it is difficult to make causal links with epidemiological studies, the growing human literature correlating environmental BPA exposure to adverse effects in humans, along with laboratory studies in many species including primates, provides increasing support that environmental BPA exposure can be harmful to humans, especially in regards to behavioral and other effects in children. (Rochester, 2013, p. 132)*

Swan (2008) argues that epidemiological studies are useful in that they can help establishing potential links, but notes that, as they cannot be used to establish the causality between exposure and effects, establishing causality requires support from animal and mechanistic studies.

> *Epidemiological studies may be the "gold standard" for assessing risks to human health, but these are expensive (costing upwards of one-million US dollars per study) and slow (typically a minimum of five years from conception of a human health study to results). Moreover, epidemiology is limited in its ability to draw causal inferences. Therefore, it must be informed by, and closely linked to, animal studies in which mechanism and causality can be demonstrated. (Swan, 2008, p.183)*

4.3.3 Epidemiological Harm Argument (EHA): Epidemiological studies can be used to establish links between exposure and health impacts through mounting correlative evidence

This argument asserts that the cumulative evidence from multiple epidemiological studies on their own are sufficient to establish links between endocrine disruptors and health impacts. Further evidence from animal studies is not required to validate such links, although such studies can lend support to conclusions based on epidemiological studies by providing causal evidence of mechanistic pathways. According to this argument, the strengths of epidemiological studies are seen as key and the limitations are less relevant.

As in EChA, the ability of epidemiological studies to measure long term exposure and subtle endpoints is emphasized in this argument. Similarly, to EChA and DCA, the EHA also asserts that epidemiological studies are more reflective of actual human exposure.

However, EHA differs from DCA by asserting that epidemiological studies are seen as having greater relevance for human effects than animal studies. EHA differs from EChA in that it goes further in asserting that epidemiological studies can be used to establish links between exposure and effect, not just to evaluate and describe the effects.

Schug et al. (2011) argues that it is already known that EDs contribute to adverse outcomes, and the correlative evidence is sufficient to assert this. They therefore hold that it is not necessary to isolate a specific chemical as the sole cause. According to this line of reasoning, it is consequently not necessary to prove that a specific ED is the sole cause of an effect for a link between exposure and effect to be asserted.

> *… it is not surprising that EDCs contribute to many adverse reproductive health outcomes in developing and adult humans. Epidemiological data has revealed an increase in male reproductive function disorders over the past 50 years, suggesting a correlative relationship with the increasing amounts of EDCs in the environment. (Schug et al., 2011, p. 211)*

In contrast to ITA, which asserts that animal studies can infer causality as they allow for the removal of other causal factors and a clear pathway between exposure and effect, the EHA assumes that all EDs contribute to these health impacts.

The EChA assumes that exposure takes place in a complex system with multiple chemicals and exposure routes. The way in which effects are manifested is also complex, with delayed exposure and subtle endpoints. The argument embraces a high tolerance of this uncertainty, which is why this argument holds that data from epidemiological studies trump data from animal studies as the former acknowledge the complexity of the system whereas the latter are held to be over-simplification. Animal studies are consequently not needed to confirm that EDs cause harmful effects as although they can demonstrate a clear pathway from exposure to effect, they cannot be used as proof of human effects. This relies on the idea of environmental exposure, that environment and health are inextricably linked. Epidemiological studies are the best way to see this link.

> *Epidemiological evidence increasingly suggests that environmental exposures early in development have a role in susceptibility to disease in later life (Jirtle & Skinner, 2007, pg. 253)*
>
> *Here we bring together the mounting evidence that environmental influences early in development are linked to disease phenotypes through modifications of the epigenome.' (Jirtle & Skinner, 2007, pg. 255)*

The second justification for EHA is that the more epidemiological studies align or pick up similar effects, even if they are from different chemicals, they all add to the burden of proof. It is the burden of epidemiological evidence, not just individual studies, that allow claims to be made that a specific effect is linked to a specific exposure. Schug et al. (2011) implies failing to find an association in an epidemiological study is not evidence for no or inconclusive effects, because of delayed effects presentation.

> *…the potential lag between exposure to EDCs and the manifestation of a clinical reproductive disorder is of critical concern. In humans, this period may be years or decades post exposure because sexual maturity and fertility cannot be assessed until the exposed individual has attained a certain age. (Schug et al., 2011, p. 211)*

Vandenberg et al. (2012) also claims that association can be used to establish links between exposure and effect, explaining that the association demonstrated in epidemiological studies provides conclusive evidence that EDCs influence certain human disorders.

> *Whether low doses of EDCs influence certain human disorders is no longer conjecture, because epidemiological studies show that environmental exposures to EDCs are associated with human diseases and disabilities…Thus, fundamental changes in in chemical testing and safety determination are needed to protect human health (Vandenberg et al., 2012, p. 378).*

4.3.4 Lack of Association Argument (LAA): Epidemiological studies have limited value because they cannot sufficiently measure association

The Lack of Association Argument (LAA) asserts that inferences made based on epidemiological studies are of limited value because they cannot sufficiently demonstrate association, or that a demonstrated association does not necessarily indicate a link between an exposure and effect. This argument asserts that certain criteria are required to demonstrate association, particularly knowing when, where and how populations are exposed to EDs. According to this argument, if these variables cannot be accurately measured, then associative claims cannot be made.

Kabir et al. (2015) asserts that certain correlative trends do not constitute proof.

> *These increases might partly reflect an increase in the likelihood of diagnosis but do not constitute proof of the impact of endocrine-disrupting chemicals. (Kabir et al., 2015, p. 252)*

The paper argues that time trend studies are not well suited to study associations because exposure cannot be controlled. Therefore, knowledge of exposure is needed to assert association.

> *Time trends and ecological studies are not well suited to study a possible association between exposure to endocrine disrupting chemicals and risk of disease, as assessment of exposure is extremely difficult (Kabir et al., 2015, p. 252)*

The second reason that underpins this argument is that epidemiological studies oversimplify to make claims of association. This reason asserts that these studies either eliminate or fail to include other potential factors that impact the effect, such as exposure burdens. As these epidemiological studies are not reflective of the whole system of exposure, then it is not possible to use the findings from these studies to claim that association is due to the factors chosen to be included in the study.

For example, Daughton & Ternes (1999) argue that these studies are not suitable for risk assessment as they oversimplify by not factoring in body burdens and exposure to other chemicals. In their paper, it is implied that epidemiology studies focusing on one chemical ignore the other contributive factors.

> *Current comprehensive environmental risk assessments and epidemiologic studies do not factor in exposures/body burdens from PPCPs and therefore may be flawed by over simplicity. (Daughton & Ternes, 1999, p. 908)*
>
> *Because of the silent and pervasive nature of environmental phthalate exposure, traditional epidemiological methods of exposure assessment (questionnaires, medical records, etc.) are of limited usefulness in*

> *determining individual exposure. Instead, biomarkers of exposure are preferred. (Swan, 2008, p. 178)*

In EHA the main limitation of epidemiological studies as asserted by the Lack of Association Argument, that they cannot provide clear links between individual chemicals and individual effects when there are multiple other factors considered, is of less importance. This is because in this context epidemiological studies are seen as relevant for proving the links between exposure to EDs collectively and their collective health impacts; of evidence that the environment and health are inextricably linked. Therefore, the evidence from multiple epidemiological studies can be used to make claims of links.

4.4 Identification of two hypothetical Ways of Reasoning in ED Literature

The identification and characterization of the arguments I have presented in chapter 4 led to the impression that certain arguments were aligned and could be combined into two distinct Ways of Reasoning (WOR), which I have chosen to call Inherent Hazard Reasoning and Collective Harm Reasoning, and that these rest on diverging framings of the problem. As further outlined below, it also appeared as if the two WOR rest on different preferences with regards to the type of studies that are held to be most suitable for assessing the potential harm, and that these WORs were tied to the type of knowledge needed to formulate policy. The hypothetical WORs and the arguments supporting them are summarized in table 2.

Table 2 Overview of how the arguments identified in the present thesis hypothetically come together as two major ways of reasoning. The table includes the colour coding used for analysis of abstracts for the presence of the arguments identified in the thesis.

WOR		Argument	#	Acronym
Inherent Hazard	Appears to argue that mechanistic and controlled animal	Inherent Toxicity	1	ITA
		Mechanistic Causality	2	MCA
		Mechanistic Extrapolation	3	MEA

	exposure studies are required to inform chemical hazard assessment for regulation.	Individual Chemical	5	ICA
		Epidemiological Characterization	10	EChA
		Dual Evidence Causality	11	DECA
		Lack of Association Argument	13	LAA
Collective Harm	Holds that health effects from chemical exposure are collectively caused by chemicals in the environment. Mechanistic studies are therefore less relevant because there is limited need to attribute effects to individual chemicals. More knowledge is required to understand the risks, exposure routes, and effects of chemical exposure so that the collective risks can be effectively mitigated. Based on this line of reasoning, this WOR criticizes regulations that deal with chemicals individually or fail to include the endpoints better characterized by epidemiological studies and latent exposure. CRA is a constituent argument of this WOR	Analogous Effect Argument	4	AEA
		Contextuality	6	CA
		Reductive Mechanistic	7	RMA
		Collective Harm	8	CHA
		Epidemiological Harm Argument	9	EHA
		Difficulty in Extrapolation	10	DEA
		Epidemiological Characterization	11	EChA
		Dual Evidence Causality	12	DECA

Below, I provide a couple of examples to illustrate, using the colour in table coding for the arguments

For example, the abstract by Gray (2000), displays the following arguments that build to the Inherent Hazard WOR: ITA, MCA, ICA . The arguments are highlighted as shown in table 2.

> *In mammals, exposure to antiandrogenic chemicals during sexual differentiation can produce malformations of the reproductive tract.*

Perinatal administration of AR antagonists like vinclozolin and procymidone or chemicals like di(2-ethylhexyl) phthalate (DEHP) that inhibit fetal testicular testosterone production demasculinize the males such that they display reduced anogenital distance (AGD), retained nipples, cleft phallus with hypospadias, undescended testes, a vaginal pouch, epididymal agenesis, and small to absent sex accessory glands as adults. In addition to DEHP, di-n-butyl (DBP) also has been shown to display antiandrogenic activity and induce malformations in male rats. In the current investigation, we examined several phthalate esters to determine if they altered sexual differentiation in an antiandrogenic manner. We hypothesized that the phthalate esters that altered testis function in the pubertal male rat would also alter testis function in the fetal male and produce malformations of androgen-dependent tissues. In this regard, we expected that benzyl butyl (BBP) and diethylhexyl (DEHP) phthalate would alter sexual differentiation, while dioctyl tere- (DOTP or DEHT), diethyl (DEP), and dimethyl (DMP) phthalate would not. We expected that the phthalate mixture diisononyl phthalate (DINP) would be weakly active due to the presence of some phthalates with a 6-7 ester group. DEHP, BBP, DINP, DEP, DMP, or DOTP were administered orally to the dam at 0.75 g/kg from gestational day (GD) 14 to postnatal day (PND) 3. None of the treatments induced overt maternal toxicity or reduced litter sizes. While only DEHP treatment reduced maternal weight gain during the entire dosing period by about 15 g, both DEHP and DINP reduced pregnancy weight gain to GD 21 by 24 g and 14 g, respectively. DEHP and BBP treatments reduced pup weight at birth (15%). Male (but not female) pups from the DEHP and BBP groups displayed shortened AGDs (about 30%) and reduced testis weights (about 35%). As infants, males in the DEHP, BBP, and DINP groups displayed femalelike areolas/nipples (87, 70, and 22% ($p < 0.01$), respectively, versus 0% in other groups). All three of the phthalate treatments that induced areolas also induced a significant incidence of reproductive malformations. The percentages of males with malformations were 82% ($p < 0.0001$) for DEHP, 84% ($p < 0.0001$) for BBP, and 7.7% ($p < 0.04$) in the DINP group. In summary, DEHP, BBP, and DINP all altered sexual differentiation, whereas DOTP, DEP, and DMP were ineffective at this dose. Whereas DEHP and BBP were of equivalent potency, DINP was about an order of magnitude less active.

An example of a Collective Risk abstract is Baillie-Hamilton, (2002) which displays the following arguments forming the Collective Risk WOR: EChaA, EHA, CA, DEC, CRA,

The number of obese people worldwide has escalated recently, revealing a complex picture of significant variations among nations and different profiles among adults and children, regions, and occupations. The *commonly held causes of obesity - overeating and inactivity - do not explain the current obesity epidemic. There is evidence of a general decrease in food consumption by humans and a significant decline in their overall levels of physical activity. There is also more evidence to indicate that the body's natural weight-control mechanisms are not functioning properly in obesity. Because the obesity epidemic occurred relatively quickly, it has been suggested that environmental causes instead of genetic factors maybe largely responsible. What has, up to now, been overlooked is that the earth's environment has changed significantly during the last few decades because of the exponential production and usage of synthetic organic and inorganic chemicals. Many of these chemicals are better known for causing weight loss at high levels of exposure but much lower concentrations of these same chemicals have powerful weight-promoting actions. This property has already been widely exploited commercially to produce growth hormones that fatten livestock and pharmaceuticals that induce weight gain in grossly underweight patients. This paper presents a hypothesis that the current level of human exposure to these chemicals may have damaged many of the body's natural weight-control mechanisms. Furthermore, it is posited here that these effects, together with a wide range of additional, possibly synergistic, factors may play a significant role in the worldwide obesity epidemic.*

These abstracts demonstrate that the arguments can appear together and provide tentative evidence for the WORs occurring in the peer reviewed literature.

4.5 Validating the identified arguments and testing the hypothesis of two distinct WOR

The follow-up study was designed to investigate if the 13 identified arguments exist in other papers. It was also designed to systematically test the hypothesis that the arguments align to form the two hypothetically outlined WOR. All arguments apart from the LAA argument were found in the abstracts with a frequency varying from 7 for ITA and RMA to 27 and 33 for CRA and MCA. This provides further support for the conclusion drawn from the initial analysis: that these arguments are commonly present in peer reviewed papers on EDs and human health.

The co-occurrence results are shown in Table 3. The darker the shade of red, the more two arguments co-occur together in the same abstract. The 13 arguments are organized such that the 7 perceived as belonging to the hypothetically formulated WOR-1 are placed above the 4 arguments perceived as belonging to the hypothetically formulated WOR-2. Notably, there is no co-occurrence pattern supporting the proposed hypothesis. In other words, the follow-up study did not provide any quantitative evidence supporting the hypothesis that the arguments in the hypothetically formulated WORs predominantly occur together. It is, however, not possible to reject the hypothesis based on the follow-up study. This is because this part of the study was based on an analysis of abstracts whilst the initial analysis considered full papers where arguments were found in different sections; notably the introductions, discussion and conclusions. Therefore, it cannot be excluded that arguments co-occur to form the proposed WORs in other parts of the papers. It also must be noted that the prescriptive coding in the validation study does not assess the strength of the arguments present, merely whether they occur or not. Further analysis of full papers is therefore needed to test the existence of the proposed WOR.

Table 3. Co-occurrence table of the constituent arguments of the two proposed WORs.

Argument	ITA	ICA	MCA	MEA	DEC	EChA	CHA	CA	RMA	EHA	DEA	AEA
ITA		4	3	1	2	3	4	6	1	2	0	0
ICA	4		4	2	1	1	4	6	0	2	0	1
MCA	3	4		7	8	8	11	16	1	13	1	9
MEA	1	2	7		3	3	6	6	0	6	1	5
DEC	2	1	8	3		8	5	8	1	8	0	2
EChA	3	1	8	3	8		7	9	1	9	0	5
CHA	4	4	11	6	5	7		14	3	10	1	6
CA	6	6	16	6	8	9	14		2	9	0	7
RMA	1	0	1	0	1	1	3	2		0	0	0
EHA	2	2	13	6	8	9	10	9	0		1	6
DEA	0	0	1	1	0	0	1	0	0	1		0
AEA	0	1	9	5	2	5	6	7	0	6	0	

Chapter 5: Discussion

In Chapter 2, I explored characteristics of epistemic perspectives based on a selected number of scholars on whom I am drawing in my research. The review shows that the weight given to different characteristics differs among the authors; for example, Haas describes epistemic communities as groups of people who are united by shared normative and principled beliefs and a value-based rationale for a common policy-based enterprise whilst Knorr Cetina, when speaking of epistemic cultures, refers to a shared valuing of knowledge. Both of these theories can be helpful when discussing the findings of the present study, as further elaborated below.

Endocrine disruptor research is an interdisciplinary field and there is a diverse representation of methods, approaches, and disciplinary backgrounds. Therefore, there is likely to be differences in knowledge creation practices, as Knorr Cetina describes in her work characterizing epistemic cultures.

Endocrine disruptor research is also a policy relevant research field and researchers commonly allude to policy in their papers. The present study suggests that the preferred policy options are tied to preferred scientific approaches, which indicates that the arguments hold positions on how knowledge can be used in policy, which is where Haas becomes helpful in shedding light on the findings.

5.1 Core differences among the ways of reasoning underlying the 13 arguments

The 13 arguments identified in the present study appear to diverge on three central points:

1. Problem approach or conceptualisation

2. The type of knowledge inferences that are appropriate
3. The type of knowledge that should be given preference in policy

5.1.1 Problem Approach or Conceptualisation

Chapter 2 highlighted how different ways of knowing can diverge based on the chosen perspective or starting point. Certain arguments appear to reason differently on a conceptual or approach basis. These arguments reason why a certain approach or conceptualization is superior to another. I am defining approach by how the problem is framed or perceived.

For example, CRA and ICA advocate for different approaches to evaluate risk: while CRA asserts that the potential harm posed by chemicals in the environment are collective and it is difficult to attribute the risks to individual chemicals, the ICA asserts that the potential harm is inherent in certain individual chemicals and can thus be evaluated at the individual chemicals level.

Several of the authors discussed in Chapter 2 discuss how the differences in starting point impact the knowledge production process. Lacey, for example, describes how the choice of methodological research strategies constrains the questions that can be asked and, consequently, the resulting findings.

The differences among the arguments can been seen as a form of system framing. Meadows describes system framing as different conceptualizations, or perceptions, of the overall system. For researchers to study any complex system, a boundary must be drawn, and some of the relevant components must be selected. This effectively simplifies the system, so it can be feasibly studied. The approaches appear to differ in where they draw the boundaries of the system, and which components they prioritize for study. For example, CRA may have a wider system boundary (by including contextual and environmental factors) than the ICA but exclude certain system components (such as individual contaminants), that are included in ICA.

Knorr Cetina theorizes the concept of epistemic strategies in their work, meaning the different approaches used to ensure and justify the validity of a result. She explains that the two disciplines that she studied, molecular biology and physics, were fundamentally using different approaches to knowledge creation. For example, physics attempts to indirectly construct reality, whilst molecular biology shows preferences for experimental evidence that does not rely on speculative interpretation.

Earlier studies have identified and characterized different approaches in the ED field. Clahsen et al., for example, found that the two papers analyzed in their study framed the endocrine system very differently; with Lamb framing it as resilient and designed to cope with environmental exposures though homeostasis and Bergman et al. framing it as vulnerable to environmental exposures which can potentially cause irreversible changes. Clahsen et al conclude that different researchers can frame the same topic fundamentally differently. In this study, CRA and ICA represent two different framings of ED risk; that risk emerges from multiple chemicals, or that risk is attributed to certain chemicals more than others.

A difference between this study and previous studies is that this study identified that preferences in approach are linked to preferences in methods. This appears to be because some methods are more compatible with the approaches than others. For example, the ICA conveys the importance of methods that evaluate chemicals individually, such as mechanistic studies and whole animal studies.

McIlroy et al., who carried out a focus group study of two groups assessed as belonging to different epistemic cultures, conclude that the one group framed endocrine disruptors as 'not necessarily having any clear distinguishing features when compared to other types of toxicants, therefore not meriting any special regulatory treatment or specific research attention as such.' In contrast, the other group framed EDs as an intractable management problem with many facets, for example, requiring consideration of both human and wildlife exposure, many possible sources of origin and complicated by mixture effects. The authors argue that framing has the potential to narrow down the possible solutions to a problem. This finding is particularly relevant for this study, because certain arguments are made up of components that rely on each other for reasoning.

5.1.2 Knowledge Inferences

For most of the analyzed papers, the conclusion they are working towards is whether a chemical/multiple chemicals are responsible for a health effect(s). The analysis suggests that the identified arguments diverge in three central aspects:

- The type of inferences that can be drawn from certain evidence
- The amount of evidence needed to draw a conclusion
- The question/s that ought to be answered

First, knowledge-based differences among the arguments may be due to inferences that are conveyed as appropriate to make, based on certain type of evidence. For example, the arguments EChA, DEC and LAA vary in the inferences they think can be made from epidemiological studies. As discussed in Chapter 4, LAA conveys that it is difficult to even make claims of association from epidemiological studies whilst EHA argues that contributive claims can be made from epidemiological studies. There is a huge disparity as one argument views these studies as not sufficient to claim association, whilst another views them as sufficient to draw conclusions about links between exposure and effect. MEA argues that extrapolation can be made through animal studies through the medium of mechanism, AEA argues that extrapolation can be made through animal studies through the medium observable effects whilst DEA argues that extrapolation is difficult through any medium.

Second, differences among the identified arguments can also be due to them requiring different amounts of evidence to draw conclusions. For example, DEC requires both epidemiological and mechanistic evidence to support a conclusion. EHA requires multiple epidemiological studies that align to some degree. LAA appears to question whether this conclusion can even be proven through these studies.

Third, the identified arguments also differ in their focus, or the type of question that are conveyed as relevant. For example, EChA sees it most relevant to explore the effects resulting from exposure, with less focus on the mechanisms underlying them. DEC seeks mechanistic evidence underpinning epidemiological findings.

All three aspects are relevant for the question discussed in the next sections, which is what type of knowledge should be considered in policy.

A knowledge inference conflict was identified by Clahsen, who found a disagreement over the application of the Bradford-Hill criteria. These criteria are sometimes used to evaluate the suitability of epidemiological studies for validating a proposed relationship between exposure and effect. Clahsen argues that the starting point of Lamb et al., (2014) was that the criteria are "adequate starting point to unequivocally establish causation" (p. 71), whilst they argue that the starting point of Bergman et al., (2015) is that the criteria "cannot be applied unequivocally and the social, economic and political context may impact the evidence required to take policy action." (p. 71). Therefore, Clahsen argues that the two authors disagree on what evidence was required for epidemiological studies to draw a reliable conclusion and justify policy action. Although this is also a policy disagreement, which will be discussed further below, this is primarily a disagreement on the standards required for valid knowledge.

If different papers take different insights from similar studies or disagree on what insights that can be drawn based on certain methods then the diverging conclusions may be irreconcilable. By this I mean unable to be resolved through a common reasoning. Further research is unlikely to resolve these differences in reasoning as a reduction in uncertainty is unlikely to change how different researchers perceive the methods. It was difficult to always recognize the inferences being made on different methods within the papers, as they were not often stated explicitly by the authors.

This does then create uncertainty for policy makers who could be unsure what the papers are implying.

5.1.3 How much knowledge is needed for policy action?

The differences in reasoning underlying the different arguments identified in the present study do not only concern preferred methods, but also how much knowledge is required for policy intervention, and the type of policy interventions that are, in parallel, advocated for.

The Bradford-Hill disagreement in Clahsen's study also has policy implications because Bergman implies that scientific certainty – meaning definite proof of a causal relationship – is not required for policy action. The CRA in this study seems to align with Bergman's view that scientific certainty is not required for action, or that certain scientific uncertainty needs to be considered in the context of how the research is applied and not be isolated from it.

Some of the arguments identified in this study stress that absence of evidence is not evidence of absence. For example, the CRA and the EHA both take a precautionary viewpoint in that current evidence is sufficient to support policy intervention, even if uncertainties remain. The EHA in particular asserts absence of evidence is not evidence of absence due to the delayed latent effects characteristic of ED exposure. Other papers stress that the uncertainty is currently too high to support intervention. These arguments assert that if the uncertainty is too high, it is unclear what the policy interventions should be. Casals-Casas & Desvergne argue that risk factors must be understood before intervention is taken; in other words, remaining uncertainty stifles meaningful action.

5.1.4 Type of knowledge needed for policy

The arguments identified in this study also diverge in the type of knowledge needed to inform policy. For example, CRA argues that chemicals should be handled collectively in policy whilst the ITA and ICA advocate for the identification of the highest risk chemicals and directing resources towards these chemicals.

Haas' work is relevant for the interpretation of these findings as he proposes that a uniting factor of epistemic communities is the way in which they use and advocate for the knowledge they produce. Certain arguments do imply how their knowledge should be used; suggesting that the differences in the ways of reasoning underlying these arguments are not just about knowledge, but policy as well, and the type of knowledge ought to be used for policy. For example, the ITA asserts that the hazards of individual chemicals should be investigated through controlled exposure and mechanistic studies. This knowledge can then inform policy risk assessments and subsequent interventions, e.g. the banning of the highest risk chemicals. The EChA asserts that more research should look to understand the risk factors and consequences of chemical exposure to help inform risk mitigation policies.

Policy processes in chemical regulation do, to a considerable extent, rely on scientific information to inform decisions. The weight of evidence (WOE) approach, for example, assumes that some evidence is worth more than other and that different types of evidence can be compared against each other. It may be seen as self-explanatory that some type of evidence is of more value than other types of evidence but the WOE approach is heavily debated and several papers address the need (and challenge) of assessing not only the reliability but also the relevance of evidence in the field of ED research and policy (Ingre-Khans et al., 2020; Kase et al., 2016; Science Advice for Policy by European Academies, 2019). Clahsen et al. discusses WOE debate in their study, as they find that the two papers express a preference for a specific WOE approach and they question whether an objective WOE approach is even possible for assessing the potential harm posed by EDs.

> *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…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 (Clahsen et al., 2020, p. 73)*

Although certain papers in this study did refer to a WOE approach, many did not specify mechanisms for achieving this. This suggests that the focus of the papers is not developing ways of knowledge integration but arguing why their own approach is superior.

CRA and ICA appear to take as their point of departure the knowledge required for the operation of the presently existing regulatory framework for assessing chemical risk, and they consequently prioritize the type of knowledge that is compatible with this approach. For example, as the policy system is based on analysis of individual chemicals, the focus should be on optimizing methods that allow for rapid individual chemical analysis. This is the link between the MCA, and ICA which forms the basis for the hypothetical inherent hazard way of reasoning. This could be interpreted as different pragmatic approaches to the regulatory system rather than epistemic differences, but this is beyond the scope of this thesis.

In contrast, arguments EChA, EHA RMA and DEA appear to set out to challenge, critique, or improve the operation of the existing regulatory framework by advocating for alternative approaches for risk analysis. The line of reasoning in these arguments is that certain approaches, methods, and endpoints are not properly considered in the current policy system. For example, the CA and the RMA both challenge the use of mechanistic studies in regulation because they say they exclude factors, over-simplify or make undue inferences from these studies.

This is contrasted by other arguments, such as ICA and the MCA, which say that regulation has to deal with a huge number of chemicals, and mechanistic studies are capable of identifying the highest risk. These arguments also asserts that regulation predominately works by regulating single chemicals. The knowledge these studies produce, this argument goes, is therefore providing the knowledge the regulation system needs.

5.2 Implications of the existence of multiple Ways of Reasoning

A lingering question is if some of the reasonings that underlie the different arguments characterized in the present study are irreconcilable or not. Below it is discussed how components of arguments diverge in the way they interpret similar types of knowledge.

An example is the difference between the MCA and the RMA. The first asserts that mechanistic evidence can be used to infer the health risk of certain chemicals, while the second asserts that mechanistic studies oversimplify in such a way that mechanistic evidence is insufficient to assess the human health risk of chemicals. These arguments are effectively flipped versions of each other, which indicates that they could be somewhat irreconcilable.

Hacking argues that there are different ways of reasoning scientific truth, illustrated with his styles of reasoning. Similar to Kuhn, he takes a historical view on the progress of science, arguing that developments in technology allowed for the emergence of new styles. Some of the methods used in ED science appear to align with Hacking's styles of reasoning. For example, Hacking lists 'statistical' and 'laboratory' as two of his styles. Statistical appears to correspond to statistical and assay testing to laboratory, two methods that are debated extensively in the arguments in this study.

Based on Hacking's assertion that claims made under different styles of reasoning cannot be compared, it is possible that epidemiological and assay studies cannot be compared or evaluated against each other as they fall under different styles of reasoning. This does echo the difference between the arguments identified in the present thesis on how different knowledge can be used and integrated together, if at all.

The present study suggests that there are disagreements related to what inferences can be made based on data from certain methods. In Hacking's work, arguments made based on epidemiological studies would be considered the same style, and therefore comparable. However, in this study there were multiple different arguments on the inferences that can be drawn based on these studies. For example, the EHA argues that combined epidemiological studies can be used to assert causality, DEC asserts that epidemiological and animal/mechanistic studies are needed to assert causality, whilst the LAA asserts that epidemiological studies are flawed in assessing association. This suggests that there could be some incommensurability within Hacking's styles of reasoning as well as in between certain styles.

Kuhn's theories on incommensurability have some relevance for this study. He argues that science conducted under different paradigms cannot be comparable to each other as there is no common measure for assessing the merit of different scientific theories. If some of the arguments were considered as belonging to different paradigms, then Kuhn's work would imply that the knowledge under each paradigm could not be compared as there is no common measure of validity. Kuhn's theory of paradigms is difficult to apply in this study because it does not take a historical lens on how ED research has changed throughout history. Kuhn's work does not refer to paradigms existing simultaneously, so it is unlikely that the different WOR would be interpreted as paradigms through his work.

These insights suggest that some of the arguments could be irreconcilable when a piece of knowledge is required to be evaluated. For example, if an epidemiological study is used to inform policy, then how would different experts agree on its relevance if one expert thinks that links between exposure and effect is sufficient for policy action and another that it is not? The study cannot be given a knowledge value in this context if these arguments are irreconcilable.

5.2.1 The Social Dimension of Science

Flek's work challenged the idea that science is independent of the social context by demonstrating that what makes good science can be subjective. Flek argues that scientific facts in part are socially constructed because they are mutually agreed by communities of experts. Therefore, it is the social consensus, not necessarily the quality of the evidence or data that determines if something becomes known as a fact. This line of reasoning has later been followed by other sociologists, most notably Bruno Laotour and Steve Woolgar in their 1979 book *Laboratory life* and Latour's *Science in Action* (1986), which clearly demonstrate the social dimension of knowledge production. The findings of the present study reinforce these findings in that different papers presented diverging inferences and conclusions as facts or assumed knowledge without acknowledging that other papers presented conflicting findings as facts.

Knorr Cetina also observed a social dimension to knowledge creation, observing how casual conversations between lab members enforced the field norms of valid knowledge and 'good' science. The findings from the present study indicate that different research communities value different types of knowledge and that this may be reinforced by the communities.

Disagreements are, for example, rarely acknowledged and in the cases they are, it is rarely from a neutral perspective – but to refute or argue against ‘the others.’

In this study certain arguments are defined by a criticism of an argument associated with the opposing way of reasoning. For example, the RMA criticizes their perceived oversimplifications in mechanistic studies and the LAA asserts that the limitations of epidemiological studies do not easily allow for associative claims to be made. Neither of these tend to acknowledge the relevance of the other arguments.

This was also found by Durant, as the two scientists overwhelmingly emphasized the perceived differences between their community and ‘the others’. To illustrate the tendency by the papers to either ignore conflicting perspectives or present them in a negative light appears to entrench the incommensurability by focusing on why they are different, rather than how they can be integrated together.

5.2.2 Science for Policy

Following from the above, it can be concluded that the present study suggests that policy relevant debates are happening in the scientific literature. The findings also indicate that the operation of chemical regulatory systems to some extent dictate the knowledge being produced in the literature because findings are being developed with policy applications in mind. In this field, there is not a simple linear progression with the ‘best’ science informing policy. This could also be exacerbated by funding being allocated preferentially towards the fields that are deemed most policy relevant. (Sarewitz, 2004) argues that science knowledge can be chosen because “they are made sensible by, particular interests and normative frameworks." (p. 385) This implies that certain scientific knowledge is not just chosen for its strength and utility, but its ability to fit into certain co-existing systems.

Hepler-Smith (2019) argues that the approach of studying chemicals individually, identified in this study as the ICA, is not solely determined by the nature of chemicals. Instead, they argue that this approach is due to legal structures, nomenclature conventions and information lists. A similar sentiment was echoed by one of the papers in this study who argue that "at a practical level it is also unclear how regulatory decisions will be amended to account for exposure to mixtures rather than to single compounds." (Casals-Casas & Desvergne, 2011). Further research on whether research aligning with certain arguments are being disproportionality funded, combined with research that assesses the impact of individual vs collective chemical policies could identify whether this idea holds true.

The value free ideal assumes that policy disputes take place at the policy level whilst the science remains uninfluenced by policy. In the case of this study, the science-policy interface clearly extends into the scientific literature. McIlroy-Young et al. (2021) suggest that differences in expert communities could lead to a manufactured consensus, where relevant epistemic perspectives are unduly excluded in policy. Findings from this study did identify points of irreconcilability, or perceived irreconcilability, between both individual arguments and the broader ways of reasoning. This suggests the benefit of further research to identify which perspectives are being included in policy and which are excluded from policy.

Chapter 6: Conclusions

This study suggests that several distinct arguments exist in the peer reviewed ED literature. Some of these predominately value methods that enable the evaluation of individual chemicals separated from contextual factors. These arguments frame the potential harm as a property mostly inherent to the chemical, meaning that mechanistic methods and controlled animal studies are the preferred way to evaluate the potential harm posed by chemicals.

Other arguments predominately value methods that can assess the potential harm of multiple chemicals. They frame the potential harm as a contextual property that is difficult to predict without knowing amounts and pathways of exposure. These arguments hold that knowledge gained from long term epidemiological studies such as subtle endpoints, latent effects and developmental exposures are key to understanding ED risk and are needed to adequately assess the potential harm. These arguments propose that this type of knowledge is not sufficiently valued in the regulatory system, in part because it appears to clash with the shorter timescale that regulatory systems operate under. These arguments advocate for an approach that manages and mitigates the potential harm of chemicals collectively.

The study also suggests that the different arguments exhibit both knowledge and policy dimensions and that regulatory realities to some degree dictate the knowledge created. The study further suggests that the knowledge production is guided by existing requirements as it places value on certain types of knowledge. These findings suggest that a major difference between the arguments is that some rely on the assumption that their research is required to provide scientific evidence that fits within the existing regulatory framework, while others challenge the knowledge considered in existing regulatory frameworks.

There appears to be different degrees of irreconcilability between the identified arguments. The knowledge dimensions of these arguments will appear not to be resolved with more research because uncertainty reduction will not mitigate the differences in the framing of methods and approaches. However, the policy dimensions of these arguments may have potential for reconciliation.

The trialed methodological approach using bibliometric clustering to source diverse papers for grounded theory analysis showed promise at characterizing arguments existing in the literature. The arguments found were nuanced and conceptual. Strategies to improve this method would be using more recent papers as the source material, expert validation of found arguments and further testing of the hypothetical Ways of Reasoning and validation of the source bibliometrics to identify if saturation has been reached.

Bibliography

Adeel, M., Song, X., Wang, Y., Francis, D., & Yang, Y. (2017). Environmental impact of estrogens on human, animal and plant life: A critical review. *Environment International*, *99*, 107–119. https://doi.org/10.1016/j.envint.2016.12.010

Aho, B. (2017). Disrupting regulation: Understanding industry engagement on endocrine-disrupting chemicals. *Science and Public Policy*, *44*(5), 698–706. https://doi.org/10.1093/scipol/scx004

Alonso-Magdalena, P., Ropero, A. B., Soriano, S., García-Arévalo, M., Ripoll, C., Fuentes, E., Quesada, I., & Nadal, Á. (2012). Bisphenol-A acts as a potent estrogen via non-classical estrogen triggered pathways. *Molecular and Cellular Endocrinology*, *355*(2), 201–207. https://doi.org/10.1016/j.mce.2011.12.012

Amaro, A. A., Esposito, A. I., Mirisola, V., Mehilli, A., Rosano, C., Noonan, D. M., Albini, A., Pfeffer, U., & Angelini, G. (2014). Endocrine Disruptor Agent Nonyl Phenol Exerts An Estrogen-like Transcriptional Activity on Estrogen Receptor Positive Breast Cancer Cells. *Current Medicinal Chemistry*, *21*(5), 630–640. https://doi.org/10.2174/09298673113209990169

Andrade, A. J. M., Grande, S. W., Talsness, C. E., Grote, K., & Chahoud, I. (2006). A dose–response study following in utero and lactational exposure to di-(2-ethylhexyl)-phthalate (DEHP): Non-monotonic dose–response and low dose effects on rat brain aromatase activity. *Toxicology*, *227*(3), 185–192. https://doi.org/10.1016/j.tox.2006.07.022

Angle, B. M., Do, R. P., Ponzi, D., Stahlhut, R. W., Drury, B. E., Nagel, S. C., Welshons, W. V., Besch-Williford, C. L., Palanza, P., Parmigiani, S., vom Saal, F. S., & Taylor, J. A. (2013). Metabolic disruption in male mice due to fetal exposure to low but not high doses of bisphenol A (BPA): Evidence for effects on body weight, food intake, adipocytes, leptin, adiponectin, insulin and glucose regulation. *Reproductive Toxicology*, *42*, 256–268. https://doi.org/10.1016/j.reprotox.2013.07.017

Autrup, H., Barile, F. A., Blaauboer, B. J., Degen, G. H., Dekant, W., Dietrich, D., Domingo, J. L., Gori, G. B., Greim, H., Hengstler, J. G., Kacew, S., Marquardt, H., Pelkonen, O., Savolainen, K., & Vermeulen, N. P. (2015). Principles of Pharmacology and Toxicology Also Govern Effects of Chemicals on the Endocrine System. *Toxicological Sciences*, *146*(1), 11–15. https://doi.org/10.1093/toxsci/kfv082

Baillie-Hamilton, P. F. (2002). Chemical Toxins: A Hypothesis to Explain the Global Obesity Epidemic. *The Journal of Alternative and Complementary Medicine*, *8*(2), 185–192. https://doi.org/10.1089/107555302317371479

Ball, R. (2018). *An introduction to bibliometrics: New development and trends*. Chandos Publishing.

Barton-Maclaren, T. S., Wade, M., Basu, N., Bayen, S., Grundy, J., Marlatt, V., Moore, R., Parent, L., Parrott, J., Grigorova, P., Pinsonnault-Cooper, J., & Langlois, V. S. (2022). Innovation in regulatory approaches for endocrine disrupting chemicals: The journey to risk assessment modernization in Canada. *Environmental Research*, *204*, 112225. https://doi.org/10.1016/j.envres.2021.112225

Bergman, Å., Becher, G., Blumberg, B., Bjerregaard, P., Bornman, R., Brandt, I., Casey, S. C., Frouin, H., Giudice, L. C., Heindel, J. J., Iguchi, T., Jobling, S., Kidd, K. A., Kortenkamp, A., Lind, P. M., Muir, D., Ochieng, R., Ropstad, E., Ross, P. S., … Zoeller, R. T. (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. https://doi.org/10.1016/j.yrtph.2015.07.026

Buter, R. K., Noyons, E. C. M., Van Mackelenbergh, M., & Laine, T. (2006). Combining concept maps and bibliometric maps: First explorations. *Scientometrics*, *66*(2), 377–387. https://doi.org/10.1007/s11192-006-0027-y

Campbell, B. (1985). Uncertainty as Symbolic Action in Disputes among Experts. In *Social Studies of Science* (Vol. 15, pp. 429–453).

Carwile, J. L., & Michels, K. B. (2011). Urinary bisphenol A and obesity: NHANES 2003–2006. *Environmental Research*, *111*(6), 825–830. https://doi.org/10.1016/j.envres.2011.05.014

Casals-Casas, C., & Desvergne, B. (2011). Endocrine Disruptors: From Endocrine to Metabolic Disruption. *Annual Review of Physiology*, *73*(1), 135–162. https://doi.org/10.1146/annurev-physiol-012110-142200

Charmaz, K. (2000). Grounded theory: Objectivist and constructivist methods. In *Handbook of qualitative research* (2nd ed., pp. 509–535).

Charmaz, K. (2014). *Constructing grounded theory* (2nd edition). Sage.

Clahsen, S. C. S., van Klaveren, H. S., Vermeire, T. G., van Kamp, I., Garssen, B., Piersma, A. H., & Lebret, E. (2020). Understanding conflicting views of endocrine disruptor experts:

A pilot study using argumentation analysis. *Journal of Risk Research*, *23*(1), 62–80. https://doi.org/10.1080/13669877.2018.1517378

Colborn, T., vom Saal, F. S., & Soto, A. M. (1993). Developmental effects of endocrine-disrupting chemicals in wildlife and humans. *Environmental Health Perspectives*, *101*(5), 378–384. https://doi.org/10.1289/ehp.93101378

Collins, H. M., & Pinch, T. J. (2001). *The golem: What you should know about science* (2. ed., reprinted). Cambridge Univ. Press.

Cousins, I. T., Vestergren, R., Wang, Z., Scheringer, M., & McLachlan, M. S. (2016). The precautionary principle and chemicals management: The example of perfluoroalkyl acids in groundwater. *Environment International*, *94*, 331–340. https://doi.org/10.1016/j.envint.2016.04.044

Daughton, C. G., & Ternes, T. A. (1999). Pharmaceuticals and personal care products in the environment: Agents of subtle change? *Environmental Health Perspectives*, *107*(suppl 6), 907–938. https://doi.org/10.1289/ehp.99107s6907

de Bruijn, E., & Gerrits, L. (2018). Epistemic Communities in Urban Self-organization: A Systematic Review and Assessment. *Journal of Planning Literature*, *33*(3), 310–328. https://doi.org/10.1177/0885412218794083

de Melo-Martín, I. de, & Intemann. (2018). *The fight against doubt: How to bridge the gap between scientists and the public*. Oxford University Press.

Dodson, R. E., Nishioka, M., Standley, L. J., Perovich, L. J., Brody, J. G., & Rudel, R. A. (2012). Endocrine Disruptors and Asthma-Associated Chemicals in Consumer Products.

Environmental Health Perspectives, *120*(7), 935–943. https://doi.org/10.1289/ehp.1104052

Durant, J. (2021). *Uncertainty and Epistemic Cultures in the Endocrine Disruptor Scientific Deliberation*. University of British Columbia.

Everitt, B. (Ed.). (2011). *Cluster analysis* (5th ed). Wiley.

Eyles, J., Robinson, K., & Elliott, S. (2009). An epistemic community comes and goes? Local and national expressions of heart health promotion in Canada. *BMC Health Services Research*, *9*(1), 35. https://doi.org/10.1186/1472-6963-9-35

Fleck, L., Trenn, T. J., Merton, R. K., & Bradley, F. (2008). *Genesis and development of a scientific fact* (Repr. 11. Aufl). Univ. of Chicago Press.

Ford, A. R., & Scott, D. N. (2017). A Critique of Risk Disclosure as the Solution for Minimizing Toxic Exposures in Pregnancy: The Case of Nail Salon Workers. *NEW SOLUTIONS: A Journal of Environmental and Occupational Health Policy*, *27*(1), 51–67. https://doi.org/10.1177/1048291117697108

Garfield, E. (1979). Is citation analysis a legitimate evaluation tool? *Scientometrics*, *1*(4), 359–375. https://doi.org/10.1007/BF02019306

Goodson, W. H., Lowe, L., Carpenter, D. O., Gilbertson, M., Manaf Ali, A., Lopez de Cerain Salsamendi, A., Lasfar, A., Carnero, A., Azqueta, A., Amedei, A., Charles, A. K., Collins, A. R., Ward, A., Salzberg, A. C., Colacci, A. M., Olsen, A.-K., Berg, A., Barclay, B. J., Zhou, B. P., … Hu, Z. (2015). Assessing the carcinogenic potential of low-dose exposures to chemical mixtures in the environment: The challenge ahead. *Carcinogenesis*, *36*(Suppl 1), S254–S296. https://doi.org/10.1093/carcin/bgv039

Gray, L. E. (2000). Perinatal Exposure to the Phthalates DEHP, BBP, and DINP, but Not DEP, DMP, or DOTP, Alters Sexual Differentiation of the Male Rat. *Toxicological Sciences*, *58*(2), 350–365. https://doi.org/10.1093/toxsci/58.2.350

Grossenbacher, P. G., & Lovelace, C. T. (2001). Mechanisms of synesthesia: Cognitive and physiological constraints. *Trends in Cognitive Sciences*, *5*(1), 36–41. https://doi.org/10.1016/S1364-6613(00)01571-0

Grün, F., & Blumberg, B. (2006). Environmental Obesogens: Organotins and Endocrine Disruption via Nuclear Receptor Signaling. *Endocrinology*, *147*(6), s50–s55. https://doi.org/10.1210/en.2005-1129

Hacking, I. (2012). 'Language, Truth and Reason' 30years later. *Studies in History and Philosophy of Science Part A*, *43*(4), 599–609. https://doi.org/10.1016/j.shpsa.2012.07.002

Hamers, T., Kamstra, J. H., Sonneveld, E., Murk, A. J., Kester, M. H. A., Andersson, P. L., Legler, J., & Brouwer, A. (2006). In Vitro Profiling of the Endocrine-Disrupting Potency of Brominated Flame Retardants. *Toxicological Sciences*, *92*(1), 157–173. https://doi.org/10.1093/toxsci/kfj187

Harrison, K., & Hoberg, G. (1994). *Risk, science, and politics: Regulating toxic substances in Canada and the United States*. McGill-Queen's University Press.

Hatch, E. E., Nelson, J. W., Stahlhut, R. W., & Webster, T. F. (2010). Association of endocrine disruptors and obesity: Perspectives from epidemiological studies. *International Journal of Andrology*, *33*(2), 324–332. https://doi.org/10.1111/j.1365-2605.2009.01035.x

Hayes, T. B., Khoury, V., Narayan, A., Nazir, M., Park, A., Brown, T., Adame, L., Chan, E., Buchholz, D., Stueve, T., & Gallipeau, S. (2010). Atrazine induces complete feminization and chemical castration in male African clawed frogs (*Xenopus laevis*). *Proceedings of the National Academy of Sciences*, *107*(10), 4612–4617. https://doi.org/10.1073/pnas.0909519107

Hepler-Smith, E. (2019). Molecular Bureaucracy: Toxicological Information and Environmental Protection. *Environmental History*, *24*(3), 534–560. https://doi.org/10.1093/envhis/emy134

Hicks, D. (1987). Limitations of Co-Citation Analysis as a Tool for Science Policy. *Social Studies of Science*, *17*, 295–316.

Ingre-Khans, E., Ågerstrand, M., Rudén, C., & Beronius, A. (2020). Improving structure and transparency in reliability evaluations of data under REACH: Suggestions for a systematic method. *Human and Ecological Risk Assessment: An International Journal*, *26*(1), 212–241. https://doi.org/10.1080/10807039.2018.1504275

Jirtle, R. L., & Skinner, M. K. (2007). Environmental epigenomics and disease susceptibility. *Nature Reviews Genetics*, *8*(4), 253–262. https://doi.org/10.1038/nrg2045

Kabir, E. R., Rahman, M. S., & Rahman, I. (2015). A review on endocrine disruptors and their possible impacts on human health. *Environmental Toxicology and Pharmacology*, *40*(1), 241–258. https://doi.org/10.1016/j.etap.2015.06.009

Karlsson, M. (2019). Chemicals Denial—A Challenge to Science and Policy. *Sustainability*, *11*(17), 4785. https://doi.org/10.3390/su11174785

Kase, R., Korkaric, M., Werner, I., & Ågerstrand, M. (2016). Criteria for Reporting and Evaluating ecotoxicity Data (CRED): Comparison and perception of the Klimisch and CRED methods for evaluating reliability and relevance of ecotoxicity studies. *Environmental Sciences Europe*, *28*(1), 7. https://doi.org/10.1186/s12302-016-0073-x

Klinge, C. M. (2015). MiRNAs regulated by estrogens, tamoxifen, and endocrine disruptors and their downstream gene targets. *Molecular and Cellular Endocrinology*, *418*, 273–297. https://doi.org/10.1016/j.mce.2015.01.035

Knorr-Cetina, K. (1999). *Epistemic cultures: How the sciences make knowledge*. Harvard University Press.

Krimsky, S. (2000). *Hormonal chaos: The scientific and social origins of the environmental endocrine hypothesis*. Johns Hopkins University Press.

Kuiper, G. G. J. M., Lemmen, J. G., Carlsson, B., Corton, J. C., Safe, S. H., van der Saag, P. T., van der Burg, B., & Gustafsson, J.-Å. (1998). Interaction of Estrogenic Chemicals and Phytoestrogens with Estrogen Receptor β. *Endocrinology*, *139*(10), 4252–4263. https://doi.org/10.1210/endo.139.10.6216

Lamb, J. C., Boffetta, P., Foster, W. G., Goodman, J. E., Hentz, K. L., Rhomberg, L. R., Staveley, J., Swaen, G., Van Der Kraak, G., & Williams, A. L. (2014). Critical comments on the WHO-UNEP State of the Science of Endocrine Disrupting Chemicals – 2012. *Regulatory Toxicology and Pharmacology*, *69*(1), 22–40. https://doi.org/10.1016/j.yrtph.2014.02.002

Lang, I. A. (2008). Association of Urinary Bisphenol A Concentration With Medical Disorders and Laboratory Abnormalities in Adults. *JAMA*, *300*(11), 1303. https://doi.org/10.1001/jama.300.11.1303

Lee, D.-H., Steffes, M., & Jacobs, D. R. (2007). Positive Associations of Serum Concentration of Polychlorinated Biphenyls or Organochlorine Pesticides with Self-Reported Arthritis, Especially Rheumatoid Type, in Women. *Environmental Health Perspectives*, *115*(6), 883–888. https://doi.org/10.1289/ehp.9887

Li, D.-K., Zhou, Z., Miao, M., He, Y., Wang, J., Ferber, J., Herrinton, L. J., Gao, E., & Yuan, W. (2011). Urine bisphenol-A (BPA) level in relation to semen quality. *Fertility and Sterility*, *95*(2), 625-630.e4. https://doi.org/10.1016/j.fertnstert.2010.09.026

Liu, R., Nelson, D. O., Hurley, S., Hertz, A., & Reynolds, P. (2015). Residential Exposure to Estrogen Disrupting Hazardous Air Pollutants and Breast Cancer Risk: The California Teachers Study. *Epidemiology*, *26*(3), 365–373. https://doi.org/10.1097/EDE.0000000000000277

MacKay, H., & Abizaid, A. (2018). A plurality of molecular targets: The receptor ecosystem for bisphenol-A (BPA). *Hormones and Behavior*, *101*, 59–67. https://doi.org/10.1016/j.yhbeh.2017.11.001

McIlroy-Young, B., Leopold, A., & Öberg, G. (2021). SCIENCE, CONSENSUS, AND ENDOCRINE-DISRUPTING CHEMICALS: RETHINKING DISAGREEMENT IN EXPERT DELIBERATIONS. *Integrated Environmental Assessment and Management*, *17*(2), 480–481. https://doi.org/10.1002/ieam.4385

Meeker, J. D. (2012). Exposure to Environmental Endocrine Disruptors and Child Development. *Archives of Pediatrics & Adolescent Medicine*, *166*(10). https://doi.org/10.1001/archpediatrics.2012.241

Michaels, D. (2008). *Doubt is their product: How industry's assault on science threatens your health*. Oxford University Press.

Mills, J., Bonner, A., & Francis, K. (2006). The Development of Constructivist Grounded Theory. *International Journal of Qualitative Methods*, *5*(1), 25–35. https://doi.org/10.1177/160940690600500103

Moore, S. C., Matthews, C. E., Ou Shu, X., Yu, K., Gail, M. H., Xu, X., Ji, B.-T., Chow, W.-H., Cai, Q., Li, H., Yang, G., Ruggieri, D., Boyd-Morin, J., Rothman, N., Hoover, R. N., Gao, Y.-T., Zheng, W., & Ziegler, R. G. (2016). Endogenous Estrogens, Estrogen Metabolites, and Breast Cancer Risk in Postmenopausal Chinese Women. *Journal of the National Cancer Institute*, *108*(10), djw103. https://doi.org/10.1093/jnci/djw103

Myers, J. P., Antoniou, M. N., Blumberg, B., Carroll, L., Colborn, T., Everett, L. G., Hansen, M., Landrigan, P. J., Lanphear, B. P., Mesnage, R., Vandenberg, L. N., vom Saal, F. S., Welshons, W. V., & Benbrook, C. M. (2016). Concerns over use of glyphosate-based herbicides and risks associated with exposures: A consensus statement. *Environmental Health*, *15*(1), 19. https://doi.org/10.1186/s12940-016-0117-0

Nayak, M., Sreedhar, D., Prabhu, S. S., & Ligade, V. S. (2021). Global Trends in Cosmetics Use-Related Adverse Effects: A Bibliometric Analysis of Literature Published during 1957–2021. *Cosmetics*, *8*(3), 75. https://doi.org/10.3390/cosmetics8030075

Nelles, J. L., Hu, W.-Y., & Prins, G. S. (2011). Estrogen action and prostate cancer. *Expert Review of Endocrinology & Metabolism*, *6*(3), 437–451. https://doi.org/10.1586/eem.11.20

Nohynek, G. J., Borgert, C. J., Dietrich, D., & Rozman, K. K. (2013). Endocrine disruption: Fact or urban legend? *Toxicology Letters*, *223*(3), 295–305. https://doi.org/10.1016/j.toxlet.2013.10.022

Onodera, N., & Yoshikane, F. (2015). Factors affecting citation rates of research articles: Factors Affecting Citation Rates of Research Articles. *Journal of the Association for Information Science and Technology*, *66*(4), 739–764. https://doi.org/10.1002/asi.23209

Oreskes, N., & Conway, E. M. (2011). *Merchants of doubt: How a handful of scientists obscured the truth on issues from tobacco smoke to global warming* (Paperback edition). Bloomsbury Press.

Patterson, A. R., Mo, X., Shapiro, A., Wernke, K. E., Archer, T. K., & Burd, C. J. (2015). Sustained Reprogramming of the Estrogen Response After Chronic Exposure to Endocrine Disruptors. *Molecular Endocrinology*, *29*(3), 384–395. https://doi.org/10.1210/me.2014-1237

Pidgeon, N. (1996). Grounded Theory: Theoretical Background. In *British Psycological Society Books*.

Ramalho, R., Adams, P., Huggard, P., & Hoare, K. (2015). Literature Review and Constructivist Grounded Theory Methodology. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, *Vol 16*, No 3 (2015). https://doi.org/10.17169/FQS-16.3.2313

Rattan, S., & Flaws, J. A. (2019). The epigenetic impacts of endocrine disruptors on female reproduction across generations†. *Biology of Reproduction*, *101*(3), 635–644. https://doi.org/10.1093/biolre/ioz081

Raun Andersen, H., Vinggaard, A. M., Høj Rasmussen, T., Gjermandsen, I. M., & Cecilie Bonefeld-Jørgensen, E. (2002). Effects of Currently Used Pesticides in Assays for Estrogenicity, Androgenicity, and Aromatase Activity in Vitro. *Toxicology and Applied Pharmacology*, *179*(1), 1–12. https://doi.org/10.1006/taap.2001.9347

Ricker, M. (2017). Letter to the Editor: About the quality and impact of scientific articles. *Scientometrics*, *111*(3), 1851–1855. https://doi.org/10.1007/s11192-017-2374-2

Rochester, J. R. (2013). Bisphenol A and human health: A review of the literature. *Reproductive Toxicology*, *42*, 132–155. https://doi.org/10.1016/j.reprotox.2013.08.008

Roy, J. (2009). Estrogen-like endocrine disrupting chemicals affecting puberty in humans—A review. *Medical Science Monitor*, *15*(6), RA137–RA145.

Rundle, A., Hoepner, L., Hassoun, A., Oberfield, S., Freyer, G., Holmes, D., Reyes, M., Quinn, J., Camann, D., Perera, F., & Whyatt, R. (2012). Association of Childhood Obesity With Maternal Exposure to Ambient Air Polycyclic Aromatic Hydrocarbons During Pregnancy. *American Journal of Epidemiology*, *175*(11), 1163–1172. https://doi.org/10.1093/aje/kwr455

Saeedi Saravi, S. S., & Dehpour, A. R. (2016). Potential role of organochlorine pesticides in the pathogenesis of neurodevelopmental, neurodegenerative, and neurobehavioral disorders: A review. *Life Sciences*, *145*, 255–264. https://doi.org/10.1016/j.lfs.2015.11.006

Sarewitz, D. (2004). How science makes environmental controversies worse. *Environmental Science & Policy*, *7*(5), 385–403. https://doi.org/10.1016/j.envsci.2004.06.001

Schug, T. T., Janesick, A., Blumberg, B., & Heindel, J. J. (2011). Endocrine disrupting chemicals and disease susceptibility. *The Journal of Steroid Biochemistry and Molecular Biology*, *127*(3–5), 204–215. https://doi.org/10.1016/j.jsbmb.2011.08.007

Science Advice for Policy by European Academies. (2019). *Making Sense of Science Under Conditions of Complexity and Uncertainty* (No. 978-3-9820301-3–5). https://doi.org/10.26356/masos

Sciortino, L. (2017). On Ian Hacking's Notion of Style of Reasoning. *Erkenntnis*, *82*(2), 243–264. https://doi.org/10.1007/s10670-016-9815-9

Scott, K. (2004). Relating Categories in Grounded Theory Analysis: Using a Conditional Relationship Guide and Reflective Coding Matrix. *The Qualitative Report*, *9*(1), 113–126.

Signorile, P. G., Spugnini, E. P., Mita, L., Mellone, P., D'Avino, A., Bianco, M., Diano, N., Caputo, L., Rea, F., Viceconte, R., Portaccio, M., Viggiano, E., Citro, G., Pierantoni, R., Sica, V., Vincenzi, B., Mita, D. G., Baldi, F., & Baldi, A. (2010). Pre-natal exposure of mice to bisphenol A elicits an endometriosis-like phenotype in female offspring. *General and Comparative Endocrinology*, *168*(3), 318–325. https://doi.org/10.1016/j.ygcen.2010.03.030

Somm, E., Schwitzgebel, V. M., Toulotte, A., Cederroth, C. R., Combescure, C., Nef, S., Aubert, M. L., & Hüppi, P. S. (2009). Perinatal Exposure to Bisphenol A Alters Early

Adipogenesis in the Rat. *Environmental Health Perspectives*, *117*(10), 1549–1555. https://doi.org/10.1289/ehp.11342

Souder, W. (2012). *On a farther shore: The life and legacy of Rachel Carson* (1st ed). Crown Publishers.

Sprague, B. L., Trentham-Dietz, A., Hedman, C. J., Wang, J., Hemming, J. D., Hampton, J. M., Buist, D. S., Aiello Bowles, E. J., Sisney, G. S., & Burnside, E. S. (2013). Circulating serum xenoestrogens and mammographic breast density. *Breast Cancer Research*, *15*(3), R45. https://doi.org/10.1186/bcr3432

Swan, S. H. (2008). Environmental phthalate exposure in relation to reproductive outcomes and other health endpoints in humans. *Environmental Research*, *108*(2), 177–184. https://doi.org/10.1016/j.envres.2008.08.007

Takaoka, S., Fujino, T., Hotta, N., Ueda, K., Hanada, M., Tajiri, M., & Inoue, Y. (2014). Signs and symptoms of methylmercury contamination in a First Nations community in Northwestern Ontario, Canada. *Science of The Total Environment*, *468–469*, 950–957. https://doi.org/10.1016/j.scitotenv.2013.09.015

Taxvig, C., Hass, U., Axelstad, M., Dalgaard, M., Boberg, J., Andeasen, H. R., & Vinggaard, A. M. (2007). Endocrine-Disrupting Activities In Vivo of the Fungicides Tebuconazole and Epoxiconazole. *Toxicological Sciences*, *100*(2), 464–473. https://doi.org/10.1093/toxsci/kfm227

Taylor, P. N., Albrecht, D., Scholz, A., Gutierrez-Buey, G., Lazarus, J. H., Dayan, C. M., & Okosieme, O. E. (2018). Global epidemiology of hyperthyroidism and hypothyroidism. *Nature Reviews Endocrinology*, *14*(5), 301–316. https://doi.org/10.1038/nrendo.2018.18

Trasande, L., Cronk, C., Durkin, M., Weiss, M., Schoeller, D. A., Gall, E. A., Hewitt, J. B., Carrel, A. L., Landrigan, P. J., & Gillman, M. W. (2009). Environment and Obesity in the National Children's Study. *Environmental Health Perspectives*, *117*(2), 159–166. https://doi.org/10.1289/ehp.11839

United Nations Environment Programme & World Health Organization. (2013). *State of the science of endocrine disrupting chemicals—2012 an assessment of the state of the science of endocrine disruptors*. WHO : UNEP. https://wedocs.unep.org/bitstream/handle/20.500.11822/12223/State%20of%20the%20Science%20of%20EDCs%20Summary%20Report%202012.pdf?sequence=1&isAllowed=y

Vandenberg, L. N. (2014). Non-Monotonic Dose Responses in Studies of Endocrine Disrupting Chemicals: Bisphenol a as a Case Study. *Dose-Response*, *12*(2), dose-response.1. https://doi.org/10.2203/dose-response.13-020.Vandenberg

Vandenberg, L. N., Colborn, T., Hayes, T. B., Heindel, J. J., Jacobs, D. R., Lee, D.-H., Shioda, T., Soto, A. M., vom Saal, F. S., Welshons, W. V., Zoeller, R. T., & Myers, J. P. (2012). Hormones and Endocrine-Disrupting Chemicals: Low-Dose Effects and Nonmonotonic Dose Responses. *Endocrine Reviews*, *33*(3), 378–455. https://doi.org/10.1210/er.2011-1050

Vandenberg, L. N., Hauser, R., Marcus, M., Olea, N., & Welshons, W. V. (2007). Human exposure to bisphenol A (BPA). *Reproductive Toxicology*, *24*(2), 139–177. https://doi.org/10.1016/j.reprotox.2007.07.010

Wen, H.-J., Sie, L., Su, P.-H., Chuang, C.-J., Chen, H.-Y., Sun, C.-W., Huang, L.-H., Hsiung, C. A., & Julie Wang, S.-L. (2017). Prenatal and childhood exposure to phthalate diesters and sex steroid hormones in 2-, 5-, 8-, and 11-year-old children: A pilot study of the Taiwan Maternal and Infant Cohort Study. *Journal of Epidemiology*, *27*(11), 516–523. https://doi.org/10.1016/j.je.2016.10.009

Yang, M., Ryu, J.-H., Jeon, R., Kang, D., & Yoo, K.-Y. (2009). Effects of bisphenol A on breast cancer and its risk factors. *Archives of Toxicology*, *83*(3), 281–285. https://doi.org/10.1007/s00204-008-0364-0

Yuan, N., Wang, L., Zhang, X., & Li, W. (2020). Bisphenol A and thyroid hormones: Bibliometric analysis of scientific publications. *Medicine*, *99*(45), e23067. https://doi.org/10.1097/MD.0000000000023067

Ziman, J. M. (2000). *Real science: What it is, and what it means*. Cambridge University Press.

Zoeller, R. T., Bergman, Å., Becher, G., Bjerregaard, P., Bornman, R., Brandt, I., Iguchi, T., Jobling, S., Kidd, K. A., Kortenkamp, A., Skakkebaek, N. E., Toppari, J., & Vandenberg, L. N. (2014). A path forward in the debate over health impacts of endocrine disrupting chemicals. *Environmental Health*, *13*(1), 118. https://doi.org/10.1186/1476-069X-13-118

Appendix: 60 papers for analysis

Red Cluster

Reference	Citation Weight	Title	Review
colborn t, 1993, environ health persp, v101, p378, doi 10.2307/3431890	267	Developmental effects of endocrine-disrupting chemicals in wildlife and humans.	Yes
kuiper ggjm, 1998, endocrinology, v139, p4252, doi 10.1210/en.139.10.4252	172	Interaction of Estrogenic Chemicals and Phytoestrogens with Estrogen Receptor β	No
toppari j, 1996, environ health persp, v104, p741, doi 10.1289/ehp.96104s4741	146	Male reproductive health and environmental xenoestrogens.	Yes
kavlock rj, 1996, environ health persp, v104, p715, doi 10.1289/ehp.96104s4715	144	Research needs for the risk assessment of health and environmental effects of endocrine disruptors: a report of the U.S. EPA-sponsored workshop.	Yes (workshop to identify research needs)
carlsen e, 1992, brit med j, v305, p609, doi 10.1136/bmj.305.6854.609	142	Evidence for decreasing quality of semen during past 50 years.	Yes
sharpe rm, 1993, lancet, v341, p1392, doi 10.1016/0140-6736(93)90953-e	138	Are oestrogens involved in falling sperm counts and disorders of the male reproductive tract?	Yes (but arguing for an effect)
skakkebaek ne, 2001, hum reprod, v16, p972, doi 10.1093/humrep/16.5.972	136	Testicular dysgenesis syndrome: an increasingly common developmental	Yes (research recommendations)

		disorder with environmental aspects: Opinion	
anway md, 2005, science, v308, p1466, doi 10.1126/science.1108190	128	Epigenetic Transgenerational Actions of Endocrine Disruptors and Male Fertility	No
soto am, 1995, environ health persp, v103, p113, doi 10.2307/3432519	128	The E-SCREEN Assay as a Tool to Identify Estrogens: An Update on Estrogenic Environmental Pollutants	No (assay)
kelce wr, 1995, nature, v375, p581, doi 10.1038/375581a0	120	Persistent DDT metabolite p,p'–DDE is a potent androgen receptor antagonist	No

Green Cluster

Reference	Citation Weight	Title	Review
vandenberg ln, 2012, endocr rev, v33, p378, doi 10.1210/er.2011-1050	290	Hormones and Endocrine-Disrupting Chemicals: Low-Dose Effects and Nonmonotonic Dose Responses	Yes (low dose and nonmono- tonicity)
vandenberg ln, 2007, reprod toxicol, v24, p139, doi 10.1016/j.reprotox.2007.07.010	263	Human exposure to bisphenol A (BPA)	Review (risk evaluation)

rochester jr, 2013, reprod toxicol, v42, p132, doi 10.1016/j.reprotox.2013.08.008	230	Bisphenol A and human health: A review of the literature	Review (association between exposure and health outcome)
calafat am, 2008, environ health persp, v116, p39, doi 10.1289/ehp.10753	220	Urinary Concentrations of Bisphenol A and 4-Nonylphenol in a Human Reference Population	No (epidemiological)
richter ca, 2007, reprod toxicol, v24, p199, doi 10.1016/j.reprotox.2007.06.004	146	In vivo effects of bisphenol A in laboratory rodent studies	Yes (justification for low dose risk)
lang ia, 2008, jama-j am med assoc, v300, p1303, doi 10.1001/jama.300.11.1303	145	Association of Urinary Bisphenol A Concentration With Medical Disorders and Laboratory Abnormalities in Adults	No (epidemiological, association between exposure and health outcome)
vandenberg ln, 2010, environ health persp, v118, p1055, doi 10.1289/ehp.0901716	144	Urinary, Circulating, and Tissue Biomonitoring Studies Indicate Widespread Exposure to Bisphenol A	Review (examined different studies come conflicting conclusions)
wetherill yb, 2007, reprod toxicol, v24, p178, doi 10.1016/j.reprotox.2007.05.010	136	In vitro molecular mechanisms of bisphenol A action	Review (mechanistic basis of BPA)
vom saal fs, 2007, reprod toxicol, v24, p131, doi 10.1016/j.reprotox.2007.07.005	130	Chapel Hill bisphenol A expert panel consensus statement: Integration of mechanisms, effects in animals and potential to impact	Consensus statement

		human health at current levels of exposure	
vandenberg ln, 2009, endocr rev, v30, p75, doi 10.1210/er.2008-0021	125	Bisphenol-A and the Great Divide: A Review of Controversies in the Field of Endocrine Disruption	Review (of BPA controversies)

Dark Blue Cluster

Reference	Citation Weight	Title	Review
diamanti-kandarakis e, 2009, endocr rev, v30, p293, doi 10.1210/er.2009-0002	421	Endocrine-Disrupting Chemicals: An Endocrine Society Scientific Statement	Yes
gore ac, 2015, endocr rev, v36, pe1, doi 10.1210/er.2015-1010	191	EDC-2: The Endocrine Society's Second Scientific Statement on Endocrine-Disrupting Chemicals	Yes
zoeller rt, 2012, endocrinology, v153, p4097, doi 10.1210/en.2012-1422	133	Endocrine-Disrupting Chemicals and Public Health Protection: A Statement of Principles from The Endocrine Society	Yes (drawing multiple studies together)
casals-casas c, 2011, annu rev physiol, v73, p135, doi 10.1146/annurev-physiol-012110-142200	99	Endocrine Disruptors: From Endocrine to Metabolic Disruption	Yes (main chemical compounds that may contribute to

			metabolic disruption)
grun f, 2006, endocrinology, v147, ps50, doi 10.1210/en.2005-1129	80	Environmental Obesogens: Organotins and Endocrine Disruption via Nuclear Receptor Signaling	No (proposed mechanistic for metabolic disruption)
livak kj, 2001, methods, v25, p402, doi 10.1006/meth.2001.1262	70	Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the $2^{-\Delta\Delta CT}$ Method	No (gene expression PCR method) DISCARDED
baillie-hamilton pf, 2002, j altern complem med, v8, p185, doi 10.1089/107555302317371479	69	Chemical Toxins: A Hypothesis to Explain the Global Obesity Epidemic	Review (metabolic, evidence for chemicals causing obesity)
woodruff tj, 2011, environ health persp, v119, p878, doi 10.1289/ehp.1002727	65	Environmental Chemicals in Pregnant Women in the United States: NHANES 2003–2004	No (exposure estimation to pregnant women)
rajapakse n, 2002, environ health persp, v110, p917, doi 10.1289/ehp.02110917	64	Combining xenoestrogens at levels below individual no-observed-effect concentrations dramatically enhances steroid hormone action.	No (concentration response)

hamers t, 2006, toxicol sci, v92, p157, doi 10.1093/toxsci/kfj187	62	In Vitro Profiling of the Endocrine-Disrupting Potency of Brominated Flame Retardants	No (toxicology profiling)

Yellow Cluster

Reference	Citation Weight	Title	Review
kolpin dw, 2002, environ sci technol, v36, p1202, doi 10.1021/es011055j	162	Pharmaceuticals, Hormones, and Other Organic Wastewater Contaminants in U.S. Streams, 1999−2000: A National Reconnaissance	No (environmental sampling)
daughton cg, 1999, environ health persp, v107, p907, doi 10.2307/3434573	92	Pharmaceuticals and personal care products in the environment: agents of subtle change?	Yes (environmental occurrence/holistic risk – not priority pollutants)
benotti mj, 2009, environ sci technol, v43, p597, doi 10.1021/es801845a	74	Pharmaceuticals and Endocrine Disrupting Compounds in U.S. Drinking Water	No (drinking water sampling)
staples ca, 1998, chemosphere, v36, p2149, doi 10.1016/s0045-6535(97)10133-3	71	A review of the environmental fate, effects, and exposures of bisphenol A	Yes

kim sd, 2007, water res, v41, p1013, doi 10.1016/j.watres.2006.06.034	67	Occurrence and removal of pharmaceuticals and endocrine disruptors in South Korean surface, drinking, and waste waters	No (Sampling)
fromme h, 2002, water res, v36, p1429, doi 10.1016/s0043-1354(01)00367-0	63	Occurrence of phthalates and bisphenol A and F in the environment	No (Sampling)
kabir er, 2015, environ toxicol phar, v40, p241, doi 10.1016/j.etap.2015.06.009	63	A review on endocrine disruptors and their possible impacts on human health	Yes (MoA)
kidd ka, 2007, p natl acad sci usa, v104, p8897, doi 10.1073/pnas.0609568104	61	Collapse of a fish population after exposure to a synthetic estrogen	No (Environmental Exposure study)
westerhoff p, 2005, environ sci technol, v39, p6649, doi 10.1021/es0484799	61	Fate of Endocrine-Disruptor, Pharmaceutical, and Personal Care Product Chemicals during Simulated Drinking Water Treatment Processes	No (water treatment test)
soni mg, 2005, food chem toxicol, v43, p985, doi 10.1016/j.fct.2005.01.020	60	Safety assessment of esters of p-hydroxybenzoic acid (parabens)	No (assay, parabens)

Purple Cluster

Reference	Citation Weight	Title	Review

swan sh, 2005, environ health persp, v113, p1056, doi 10.1289/ehp.8100	147	Decrease in Anogenital Distance among Male Infants with Prenatal Phthalate Exposure	Yes (exposure/association)
schug tt, 2011, j steroid biochem, v127, p204, doi 10.1016/j.jsbmb.2011.08.007	108	Endocrine disrupting chemicals and disease susceptibility	Review (disease endpoints, hormone mechanisms)
silva mj, 2004, environ health persp, v112, p331, doi 10.1289/ehp.6723	89	Urinary levels of seven phthalate metabolites in the U.S. population from the National Health and Nutrition Examination Survey (NHANES) 1999-2000.	No (exposure to phthalate)
heudorf u, 2007, int j hyg envir heal, v210, p623, doi 10.1016/j.ijheh.2007.07.011	87	Phthalates: Toxicology and exposure	Review (risk assessment)
rudel ra, 2003, environ sci technol, v37, p4543, doi 10.1021/es0264596	82	Phthalates, Alkylphenols, Pesticides, Polybrominated Diphenyl Ethers, and Other Endocrine-Disrupting Compounds in Indoor Air and Dust	No (Sampling Indoor)
gray le, 2000, toxicol sci, v58, p350, doi 10.1093/toxsci/58.2.350	81	Perinatal Exposure to the Phthalates DEHP, BBP, and DINP, but Not DEP, DMP, or DOTP, Alters Sexual Differentiation of the Male Rat	No (animal exposure test)

main km, 2006, environ health persp, v114, p270, doi 10.1289/ehp.8075	75	Human Breast Milk Contamination with Phthalates and Alterations of Endogenous Reproductive Hormones in Infants Three Months of Age	No (exposure association)
swan sh, 2008, environ res, v108, p177, doi 10.1016/j.envres.2008.08.007	73	Environmental phthalate exposure in relation to reproductive outcomes and other health endpoints in humans	No (exposure, endpoints)
hauser r, 2005, occup environ med, v62, doi 10.1136/oem.2004.017590	71	PHTHALATES AND HUMAN HEALTH	Yes (use, metabolism and exposure/association)
wormuth m, 2006, risk anal, v26, p803, doi 10.1111/j.1539-6924.2006.00770.x	71	What Are the Sources of Exposure to Eight Frequently Used Phthalic Acid Esters in Europeans?	No (exposure scenario)

Light Blue Cluster

Reference	Citation Weight	Title	Review
manikkam m, 2013, plos one, v8, doi 10.1371/journal.pone.0055387	69	Plastics Derived Endocrine Disruptors (BPA, DEHP and DBP) Induce Epigenetic Transgenerational Inheritance of Obesity, Reproductive Disease and Sperm Epimutations	No (animal test, epigenetics)

jirtle rl, 2007, nat rev genet, v8, p253, doi 10.1038/nrg2045	44	Environmental epigenomics and disease susceptibility	Yes (epigenetics)
anway md, 2006, endocrinology, v147, p5515, doi 10.1210/en.2006-0640	34	Endocrine Disruptor Vinclozolin Induced Epigenetic Transgenerational Adult-Onset Disease	
li sf, 1997, cancer res, v57, p4356	30		
manikkam m, 2012, plos one, v7, doi 10.1371/journal.pone.0031901	30	Transgenerational Actions of Environmental Compounds on Reproductive Disease and Identification of Epigenetic Biomarkers of Ancestral Exposures	No (animal test, epigenetics)
nilsson e, 2012, plos one, v7, doi 10.1371/journal.pone.0036129	27	Environmentally Induced Epigenetic Transgenerational Inheritance of Ovarian Disease	No (animal test, epigenetics)
crews d, 2007, p natl acad sci usa, v104, p5942, doi 10.1073/pnas.0610410104	26	Transgenerational epigenetic imprints on mate preference	No (animal test, epigenetics)
skinner mk, 2013, bmc med, v11, doi 10.1186/1741-7015-11-228	26	Ancestral dichlorodiphenyltrichloroethane (DDT) exposure promotes epigenetic transgenerational inheritance of obesity	No (animal test, epigenetics)

anway md, 2006, j androl, v27, p868, doi 10.2164/jandrol.106.000349	25	Transgenerational Effect of the Endocrine Disruptor Vinclozolin on Male Spermatogenesis	No (animal test, development)
wolstenholme jt, 2012, endocrinology, v153, p3828, doi 10.1210/en.2012-1195	25	Gestational Exposure to Bisphenol A Produces Transgenerational Changes in Behaviors and Gene Expression	No (animal test, epigenetics)