Technische Universität München

TUM School of Management

The determination of food safety as a societal challenge

Christine Haßauer

Vollständiger Abdruck der von der TUM School of Management der Technischen Universität München zur Erlangung des akademischen Grades einer Doktorin der Wirtschafts- und Sozialwissenschaften (Dr. rer. pol.) genehmigten Dissertation.

Vorsitzende:	Prof. Dr. Ruth Müller
Prüferinnen der Dissertation:	1. Prof. Dr. Jutta Roosen
	2. Prof. Dr. Karin Zachmann

Die Dissertation wurde am 21.02.222 bei der Technischen Universität München eingereicht und durch die TUM School of Management am 15.07.2022 angenommen.

Acknowledgements

First, I thank Prof. Roosen for the supervision of my doctoral thesis, for her support, and for providing me the opportunity work in this interesting project. Thanks to Prof. Zachmann for second examination and thank you to Prof. Müller for heading the dissertation commission.

Further thanks go to my colleagues at the Chair of Marketing and Consumer Research for being great friends, support, and critical discussion partners.

This dissertation is based on research supported by funding from the Deutsche Forschungsgemeinschaft. Thus, I thank the DFG for funding and the whole research group 2248 for creating in a productive, supportive, and enriching environment.

I thank my family and friends for their unconditional support in every stage of this project.

Summary

Food safety is a topic of great societal relevance because it affects food production, trade of food products, development, and consumer health. Different interests, potential goal conflicts, or trends like the globalization of trade and moralization of consumption require complex negotiation processes in food safety determination. Public and scientific discourses on adequate food safety levels prove that it is a complex and difficult task to define what is safe food.

In four essays, this dissertation analyzes different evidence practices from consumers and science-based risk analysis to investigate the basis of evidence in the determination of food safety. It aims to examine the related processes of evidence determination for food safety and to investigate interactions between risk analysis and consumers in these processes. The dissertation uses different qualitative approaches, including literature-based framework building and interview studies. It focusses on the European regulatory system.

The dissertation reveals complex, independent evidence practices and their interactions, and thus, the negotiation processes for food safety evidence. Additionally, it provides insights in interaction mechanisms in the integration of consumer evidence in risk analysis. It identifies strengths and weaknesses of current regulatory practice and establishes a new perspective on consumers in this process. Results explain why it is not possible to determine one universally accepted level of safe food.

This dissertation contributes to the understanding of what is societally accepted knowledge for food safety and how it is constructed by different actors. Thus, it offers a new perspective on food safety determination in form of evidence practices. The analysis of the different forms of construction of food safety evidence might increase mutual understanding, help to explain differences between societal groups, and provide approaches for an improvement of risk analysis towards an optimized integration of societal needs.

List of Contents

Acknowledgements ..... **I**

Summary ..... **II**

List of Contents ..... **III**

List of Tables ..... **VI**

List of Figures ..... **VII**

List of Appendices ..... **VIII**

List of Abbreviations and Acronyms ..... **IX**

1. Introduction **1**

1.1. Food safety determination within societal expectations 1

1.2. Research context 5

1.3. Research aim and framework 6

1.3.1. Research aim 6

1.3.2. Research framework 8

1.3.3. Structure 10

2. Background **11**

2.1. Evidence production for food safety 11

2.1.1. Risk analysis as regulatory framework for food safety 11

2.1.2. Consumers and intuitive toxicology 15

2.2. Focus areas of the dissertation 18

2.2.1. Study 1 – Conceptualizing food safety – a multidisciplinary perspective on food safety criteria 19

2.2.2. Study 2 – Food safety criteria used by consumers – a qualitative analysis of in-depth interviews 20

2.2.3. Study 3 – Practice of food safety determination – The integration of stakeholder evidence in risk assessment 21

2.2.4. Study 4 – Cost-benefit analysis as decision tool in evidence-based policy 22

2.2.5. Case studies 23

2.3. Research approach 28

3. Essays **29**

3.1. Toward a conceptual framework for food safety criteria: analyzing evidence practices using the case of plant protection products 29
3.1.1. Abstract 29
3.1.2. Introduction 30
3.1.3. Procedure 32
3.1.4. Results 36
3.1.5. Discussion 62
3.1.6. Conclusion 66
3.1.7. Acknowledgements 67
3.2. Which criteria do consumers use to evaluate the safety of food? 68
3.2.1. Abstract 68
3.2.2. Introduction 68
3.2.3. Methodology 70
3.2.4. Results 72
3.2.5. Discussion 75
3.2.6. Conclusion 77
3.2.7. Acknowledgements 77
3.3. Stakeholder Integration in the Determination of Evidence for Food Safety—Insights from European Risk Assessment Organizations 78
3.3.1. Abstract 78
3.3.2. Introduction 79
3.3.3. Background 80
3.3.4. Material and methods 84
3.3.5. Results 89
3.3.6. Discussion 99
3.3.7. Conclusions 106
3.3.8. Acknowledgements 107
3.4. Bewerten und Gewichten: Evidenz als Entscheidungshilfe in der Gesundheits- und Umweltpolitik (English translation: Evaluating and weighting: Evidence as decicion tool in health and environmental policy) 108
3.4.1. Abstract 108
3.4.2. Einleitung 108
3.4.3. Bewerten und Gewichten als Evidenzpraktiken zur Begründung von Entscheidungen 111
3.4.4. Die Etablierung der Kosten-Nutzen-Analyse in komplexen Entscheidungsprozessen 114
3.4.5. Evidenz in der Roll-Back-Malaria-Initiative 117
3.4.6. Evidenz als Entscheidungshilfe in der Kritik 124
3.4.7. Schlussfolgerungen 128
3.4.8. Acknowledgements 129

4. Conclusion 130
4.1. Main findings and discussion 130
4.1.1. The trading zone concept as research framework 134
4.1.2. Reflection on Quality 136
4.2. Implications for future research 137
4.3. Implications for risk analysis 139
4.3.1. Risk assessment 140
4.3.2. Risk management 141
4.3.3. Risk communication 142
5. References 144
6. Appendices 158

List of Tables

Table 1: Research questions and contributions of study 1 20
Table 2: Research questions and contributions of study 2 21
Table 3: Research questions and contributions of study 3 22
Table 4: Research questions and contributions of study 4 23
Table 5: Differences in risk assessment of Microbiological hazards and Chemical hazards 25
Table 6: Issues in data 41
Table 7: Preliminary forms of knowledge criteria in different steps of risk assessments 43
Table 8: Preliminary forms of knowledge criteria in different steps of hazard assessments 43
Table 9: Evaluation methods for PPPs 45
Table 10: Psychological factors in the explanation of risk perception 56
Table 11: Sample characteristics 71
Table 12: Detailed descriptions of food safety criteria 74
Table 13: Overview of interviewees 86

List of Figures

Figure 1: Actor network in evidence production and use for food safety 8
Figure 2: The practice of risk analysis 11
Figure 3: Responsibilities in EU food safety determination at the example of PPPs. . 13
Figure 4: Search strategy of the literature review 35
Figure 5: A conceptual framework for food safety criteria 37
Figure 6: Food safety criteria in the evidence practice of risk analysis 39
Figure 7: Food safety criteria in consumer evaluation 51
Figure 8: Classification of food-safety criteria 73
Figure 9: Content structure interview guide 87
Figure 10: Structure of analysis 88
Figure 11: Theoretical description of the cases. 91
Figure 12: Input: Interaction with stakeholder groups 92
Figure 13: Processing: Interactions with stakeholder groups 94
Figure 14: Output: Interactions with stakeholder groups 97

List of Appendices

Appendix A to Section 3.1. Guiding Theory 158
Appendix B to Section 3.1. Overview literature 159
Appendix C to Section 3.2. Interview guide 160
Appendix D to Section 3.2. Detailed presentation of themes and sub-themes 163

List of Abbreviations and Acronyms

ADI	Acceptable Daily Intake
ARfD	Acute Reference Dose
BfR	German Federal Institute for Risk Assessment
BVL	German Federal Office of Consumer Protection and Food Safety
BSE	Bovine Spongiform Encephalopathy
EFSA	European Food Safety Authority
EU	European Union
HACCP	Hazard Analysis Critical Control Point
HIV	Human Immunodeficiency Virus
MRL	Maximum Residue Levels
NGO	Non-Governmental Organization
NOEL	No Observed Effect Level
NOAEL	No Observed Adverse Effect Level
NRC	National Research Council
OECD	Organization for Economic Co-operation and Development
PPP	Plant Protection Product
RBM	Roll-Back-Malaria-Initiative
SPS	Agreement on Sanitary and Phytosanitary Measures
USA	United States of America
WHO	World Health Organization
WTO	World Trade Organization

1. Introduction

1.1. Food safety determination within societal expectations

Food safety standards need to fulfill a range of societal expectations. They should protect consumers from hazards occurring in food products and guarantee safe consumption. Unsafe food can cause different acute and chronic diseases and even result in death. For 2010, the World Health Organization (WHO) estimated 600 million cases of illness (European Union (EU): 23 million) and 33 million lost healthy life years (EU: 400,000) due to contaminated food (WHO, 2015). Besides the direct effects of illness, unsafe food can have different indirect social and economic consequences for consumers, for example monetary or social losses (Yeung & Morris, 2001). The access to safe food is also defined as one of the United Nations' development goals (UN, 2020; WHO, 2015). Additionally, adequate levels of food safety regulations are important for profitable food production and functioning markets. However, stricter regulations might function as non-tariff trade barriers and threaten free trade. Different levels of safety standards, for example maximum residue levels (MRLs), can hinder trade, especially for developing countries with lower standards (Disdier et al., 2008; Essaji, 2008; Henson & Loader, 2001). The harmonization of regulations is also stipulated as objective of the General Food Law of the EU (EC & EP, 2002). Although the motivations for food safety determination seem clear, it is the product of a complex societal process which entails different issues. Problems arise from different sources and affect different actors:

First, the provision of absolutely safe food is unrealistic if food safety is defined as "absence of harmful substances" (Herges et al., 2017). European food products are rated as relatively safe compared to those in developing countries. However, monitoring reports show for example stable numbers of cases of the zoonosis Campylobacteriosis with over 100,000 cases in the EU in 2017 (EFSA & ECDC, 2018). For pesticides, monitoring reports that 2,7 % of samples exceeded the MRLs in the EU in 2018 (EFSA et al., 2020). Zero risk from foods is not feasible for different reasons: Practically, in the elimination of microbiological hazards there are restrictions of resources, complex infection chains, or dynamic pathogen behavior with large amounts of scientific uncertainties (Ruzante et

al., 2010; Zwietering, 2015). Plant Protection Products[1] (PPPs) or other food technologies like irradiation may produce hazardous substances but are necessary to increase food security, or the quality of food products. Thus, a ban of these technologies is not a realistic scenario, even if they create new hazards. Conceptually, technical or chemical inventions in the food chain entail trade-offs between risks and benefits to different affected stakeholders which makes it difficult to define what is safe or not (Fischhoff & Hope, 1984; Henson & Traill, 1993). For societies, the level of safety is also an economic decision. Food safety entails costs for societies from diseases, costs for prevention measures, or costs for monitoring strategies (Focker & van der Fels-Klerx, 2020).

Second, regulations in general and specifically food safety determination should be based on sound scientific knowledge to avoid arbitrariness. This principle is stipulated in the Agreement on Sanitary and Phytosanitary Measures (SPS agreement) of the World Trade Organization (WTO) (WTO, 1995). Since then, it was implemented consistently in international trade law and food safety policy, for example, in the General Food Law in the EU (EC & EP, 2002; Wagner, 2016). Practically, the principle has partially proven to be problematic. It is debated which amount of science is considered as sufficient or when it is appropriate to take precautionary, protective measures due to scientific uncertainty (Kerr, 2009; Wagner, 2016). An example is the case of European precautionary bans of hormone beef from the Unite States of America (USA) which resulted in a longstanding legal dispute about precaution and sufficiency of scientific evidence (Ansell & Vogel, 2006). Additionally, the principle is difficult to apply in cases of scientific disagreement. The Glyphosate debate[2] in the last years in the EU exemplifies the problems of determining sufficient scientific knowledge for safety (Morvillo, 2020; Portier et al., 2016).

Third, consumers are important actors in food safety determination but are difficult to integrate in the process. They are responsible members of the food chain who contribute

[1] In this dissertation, the focus lies on plant protection products and not on the whole spectrum of pesticides. Pesticides include, for example, biocides which fall under a different regulatory framework. Nevertheless, the term pesticides is used in some contexts, depending on the focus of the original source.

[2] In 2017, the EU commission decided to extend the authorization of the herbicide Glyphosate for five more years. The process was and is accompanied by public protest, criticism by different groups, and media attention. A major point was that international scientific risk assessment organizations came to different conclusions about carcinogenicity (EC, 2021; Guyton et al., 2015; Portier et al., 2016).

to food safety actively (van der Meulen & van der Velde, 2004). Additionally, consumers are affected by food safety levels in form of food prices and availability of products (Manfreda & De Cesare, 2014). Nevertheless, it is difficult to include their perceptions into the technical determination process for food safety. It is well known that consumer risk perception differs from results of technical risk assessments. For example, consumers often underestimate natural risks such as the risks of bacteria and overestimate technological risks (Kaptan et al., 2017; Roosen et al., 2005). In the EU, 39 % of citizens reported concerns over pesticide residues in food in 2019. Nearly half of the respondents (43 %) stated that food is full of harmful substances. Contrary, the European Food Safety Authority (EFSA) describes the MRL exceedance and risk for consumers as rather low (EFSA, 2016b, 2019a; EFSA et al., 2020) and the EU regulatory system for PPPs is known as one of the strictest in the world (Handford et al., 2015). Since the 1970s, scholars provided a number of explanations for these phenomena, for example the role of risk-benefit perceptions, hazard characteristics, heuristics, values, or more recently, political orientation (Hansen et al., 2003; Kahan, 2016a; Slovic, 1987; Starr, 1969; Sunstein, 2002). Consumers talk differently about safety, have different ways to express risk perception, and internalize a different set of factors compared to scientific risk assessment. This can lead to fundamental misunderstandings and entails the risk of over- or underestimating consumer concerns (Fraiberg & Trebilcock, 1998; Leiss & Powell, 2004).

Those misunderstandings are important to avoid because consumers are an economically significant factor in food markets. Changes in their consumption behavior can have serious impact as soon as they see themselves at risk. For example, beef sales decreased by 11 % after the Bovine Spongiform Encephalopathy (BSE) crisis in the EU and even affected beef purchase behavior in the US significantly. In total, it is estimated that the BSE crises caused costs of about 50 billion US-Dollar in form of producer subsidies for the EU and lead to abnormal price drops in international commodity markets (Leiss & Powell, 2004; Schlenker & Villas-Boas, 2009). Contrary, consumer skepticism can thwart development and adoption of new and potentially beneficial technologies in food production as is shown in public resistance against the introduction of genetically modified organisms in EU agriculture. It is a well-known phenomenon that consumers

state concerns but do not buy according to these concerns which is known as the attitude behavior gap (Lusk et al., 2014).

Fourth, regulation and the level of food safety determination affects food producers. It determines how negative externalities are internalized through liability and responsibility. The General Food Law of the EU places responsibility on producers in all stages of the supply chain (Trumbull, 2006). For producers, it can be challenging to comply with food safety standards. They act on international markets in complex supply chains which can increase, for example, the risk of outbreaks of foodborne diseases. For food companies, safety incidences can cause large economic damage: Costs may occur for prevention measures or damage compensation. Costs for companies in the US due to food safety incidents were estimated up to 9 billion US Dollars per year (Hussain & Dawson, 2013). Further costs emerge from the implementation or the regular compliance with food safety standards. In the US, these costs for large food companies were estimated with $686 million implementation costs and $319 million annual costs (FDA, 2012, cited in Hessing et al., 2015). These costs may include, for example, the reformulation of products or changes in production processes (Manfreda & De Cesare, 2014).

The previous issues are amplified by recent societal drivers. Innovations in transportation, processing, and storage changed possibilities of food production and trade constantly (Lusk et al., 2014). Globalization influences food safety determination through stronger relevance of international trade flows, international trade laws, or the import of potentially dangerous products in domestic markets, which entail new challenges for traceability and responsibility allocations (Lineback et al., 2009; Trumbull, 2006). Many net importing countries rely on the import of raw materials and food products and thus on harmonized standards and control procedures (Manfreda & De Cesare, 2014). Globalized supply chains require different forms of institutions and increase the importance of supranational actions and principles like the SPS agreement (Hoffmann, 2010).

Contrary to these trends in the provision of food, on the consumer side, moralization seems to gain in importance. Traditional, cultural forms of food handling have a decreasing importance in the evaluation of food safety. Sensorial perception and cultural mechanisms are not useful to evaluate safety of the increasingly technically and industrialized produced food, for example the evaluation of chemical residues in food. This drives consumers to switch to alternative indicators such as moral beliefs (Lusk,

2013; Lusk et al., 2014; Pollan, 2006). Food choices include various moral and ethical considerations and can pose a contrasting trend to globalization (Nygard & Storstad, 1998).

1.2. Research context

The preceding arguments underline the relevance and difficulties of an optimal determination of food safety levels for all actors involved. Regulations influence health and trade and contribute to welfare and wellbeing of citizens. Costs from non-optimal food safety determination can cause large economic damage for individuals, companies, and societies. Additionally, risks become increasingly transnational and intergenerational, for example spreading of pathogens through international food chains or the long-term accumulation of chemicals in soils (Driesen, 2004). It is a complicated task to integrate societal expectations in safety determination processes. Various researchers dealt with these questions in the past decades, especially since social sciences started to point out the relevance of society in the determination of risks (for example Finn & Louviere, 1992; Pollak, 1998; Salanié & Treich, 2009).

Particularly consumers are important stakeholders in the food chain but their integration into food safety determination is problematic. Currently, the public is mostly seen as perceiving element in this process. Research focusses on the determination of relevant variables for the prediction and explanation of risk perception, for example hazard characteristics or differences in individual perception (Hassauer & Roosen, 2020). So far, risk perception is not a central element in EU food safety regulations except risk communication.

Regulations are determined evidence-based and based on the risk analysis principle which contains the steps of risk assessment, risk management, and risk communication (EC & EP, 2002). In the regulatory context, relevant evidence consists mostly of quantitative knowledge from natural sciences and not from social sciences (König et al., 2010; Zachmann, 2014). On the one hand that provides an objective and transparent basis for risk evaluations. On the other hand, standardized procedures might not be able to integrate all societal components of a complex risk issue. This might be problematic because the definition of acceptable risk and thus safety is a fundamental societal question (Fischhoff,

1994). Therefore, it is challenging to determine what evidence is seen as sufficient and if and how consumers are integrated into that process.

Thus, it is a promising approach to focus on evidence practices. Evidence itself is difficult to define. Therefore, in this dissertation it is understood as a construct of socially accepted knowledge (Cartwright, 2006; Kelly, 2016). Evidence practices are understood as procedures of evidence generation (practicing evidence) and processes of using and embedding evidence (evidencing practice) (Zachmann & Ehlers, 2019). Evidence practices are processes of producing and using knowledge. In the context of this dissertation, they are helpful to understand what safety means for different actors and how these meanings interact. The analysis of evidence practices provides an interesting approach to investigate the societal negotiation process in the determination of food safety. Additionally, it contributes to the understanding of what is societally accepted knowledge for food safety and how it is constructed by different actors. This can support increased mutual understanding, shows deficits and inconsistencies in current risk analysis processes, and points to solutions for more societally stable assessments and regulations. The next section specifies these research objectives and the framework for this dissertation.

1.3. Research aim and framework

1.3.1. Research aim

This dissertation embeds the scientific process of the determination of evidence for food safety in society. It focuses on the role of scientific risk assessment institutions and consumers in evidence determination, and specific zones of contact (trading zones) between these two publics, scientists, and consumers. Superordinate questions are when and how knowledge for safety is considered as evident (practicing evidence), and furthermore which evidence is transferred into the risk analysis process (evidencing practice). The related objectives are to examine (1) processes of evidence determination for food safety and (2) to uncover areas of co-production of evidence for food safety between scientific and consumer publics.

The first objective examines the socio-epistemic practice of determining sufficient knowledge for safety, first in risk analysis and second by consumers. It assumes that it is

not possible to define one objective level of safety. Group-specific or individual evidence practices determine whether knowledge is deemed sufficient for safety and define specific food safety criteria. That is assumed to be the origin for various controversies in the determination of food safety by producers, legislators, and consumers. Contrary, most existing literature concludes that understandings of safety differ because presumably objective knowledge is perceived differently by different groups, for example depending on political orientation (Kahan, 2016). This dissertation aims to provide a new perspective. It is based on the hypothesis that evidence for food safety is produced in different zones and contexts. Therefore, the first research question is:

A) How does the determination of sufficient evidence for food safety differ between science-based risk analysis and intuition-driven consumers?

If evidence for food safety is determined in a societal negotiation process, it is likely that different understandings from different groups overlap. Consumers have limited possibilities to influence the (mostly) scientific and highly technical practice of evidence determination in food safety but are able to put pressure on the system in form of verbal protest or purchase behavior. Additionally, they are objects of interest because their behavior influences food safety. Risk analysis needs to deal with these informal influences. Additionally, risk analysis needs to cover all relevant aspects of a risk for a society. Therefore, the second research question is

B) How do these evidence practices interact in the determination of food safety and influence the selection of evidence for food safety?

The preceding reasoning indicates that there is no simple answer on what is safe food and which evidence is sufficient to determine it. In four studies, this dissertation aims to approach this fundamental question. The four studies offer insights from different perspectives, first a conceptual perspective based on literature on risk analysis and consumers, second a perspective on the understanding of safety by consumers, third a perspective on the evidence practice in risk assessment organizations, and fourth a perspective on the risk management practice in health- and environmental policy. A more detailed description of the specific research questions is provided in chapter 2.

The focus on evidence practices allows a new perspective on the societal determination process of food safety which differs from existing approaches. It offers a possibility to

observe the encounter of two different understandings of safety – science-based risk analysis versus consumer-related intuition. This can contribute significantly to the understanding of issues, mechanisms, and interactions in food safety determination.

1.3.2. Research framework

The framework for this dissertation is based on Galisons' trading zones concept. His original concept describes how different (scientific) groups interact and influence each other's practices in so called trading zones. In these zones, for example new practices, or languages are emerging. He also describes these trading zones as possibly existing between experts of a certain field and a general public (Galison, 2010). Therefore, this concept is an interesting approach to describe and analyze the evidence practices for food safety determination. A highly technical risk analysis process meets an intuitively acting public in the practice of food production. Figure 1 illustrates the actors' network in the evidence production for food safety. Science and consumers are the main actors who interact in a trading zone: the practice of food production – including regulations, control, and production processes. These actors are embedded in wider societal structures: Societal stakeholders like NGOs initiate discourse, or international organizations like the WTO establish internationally valid regulatory frameworks. The studies of this dissertation are based on the actor's network and shed light on the dynamics within the network while considering the existence of framework conditions and external actors.

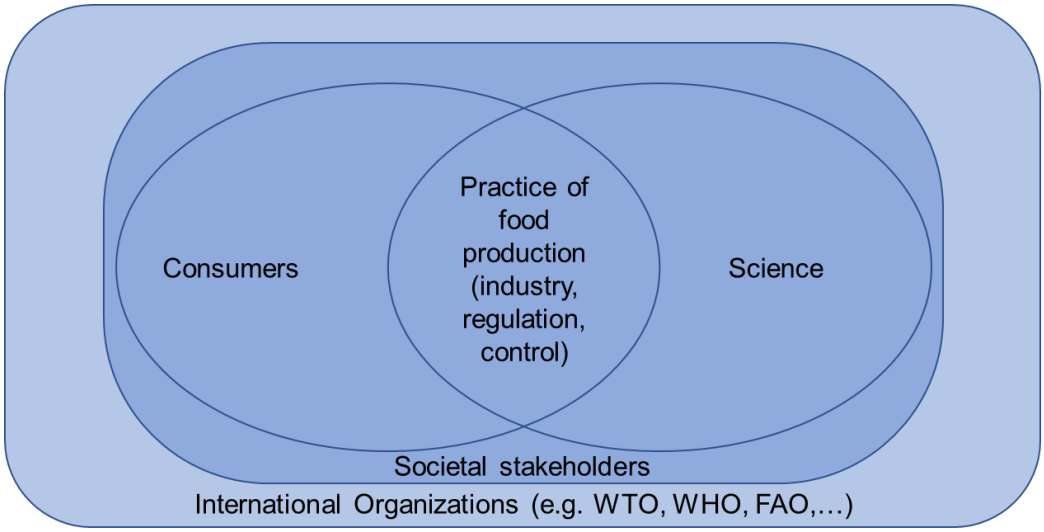


Figure 1: Actor network in evidence production and use for food safety

Expanding on Galison's trading zone concept (Galison, 2010), a central hypothesis of the dissertation is that in food safety, there are two publics which are influencing and influenced (by) each other: A scientific and a consumer public. In the trading zone, food safety is deliberated, codified, and implemented (see Figure 1). It is of special interest for the dissertation, which types of evidence are brought in from both sides and how these differences influence the practice of food safety determination. The trading zones concept has three characteristics which are important for the new perspective on food safety: First, it assumes that the two actors are not structured hierarchically but that they exist in parallel and potentially overlap. Second, it acknowledges that knowledge does not have to have the same meaning and relevance among a different group of actors. Third, there is often an imbalance of power between the actors (Galison, 2010). For food safety determination, that means that consumers are not necessarily perceivers of risk analysis, that consumers not necessarily share the meaning of knowledge from the natural sciences, and that there is a power imbalance between risk analysis and risk perception in the determination of food safety.

The studies of this dissertation are located in different areas of the concept: The first study provides a holistic view and collects criteria used by science-based risk analysis and consumers to describe safety. The second study is located on the consumer side and focusses on consumer food safety criteria. The third paper describes how risk assessors interact with stakeholders (consumers and food producers) in their evidence practice. The fourth study examines the overlapping zone and provides insights in the evidence practices in risk management.

Based on the claim "Practicing Evidence – Evidencing Practice"[3], the dissertation deals with evidence-based determination of food safety and analyzes the processes and interaction of heterogeneous evidence practices.

[3] "Practicing evidence – evidencing practice" is the overarching theme of the work of the DFG-research group 2448 of which this research project was a part of.

1.3.3. Structure

This dissertation is divided in four chapters. After introducing the work and laying out the relevance and objectives of this work, chapter two begins by providing context on the theoretical dimensions in the determination of food safety. Additionally, chapter 2 identifies the focus areas and provides the basic methodological concept for the studies. Chapter 3 presents the four studies of this dissertation. In chapter 4, results of the studies are discussed with a focus on answering the research questions and the related theoretical and practical implications.

2. Background

The following chapter outlines the conceptual and the methodological background for the studies. Section 2.1 provides background information to different processes of evidence production in food safety determination. The structure is framed within the trading zone concept which includes science-based risk analysis and consumer evaluation as relevant processes in food safety determination: First, the section lays out the role and structure of current risk analysis processes. Second, it summarizes the state of knowledge on food safety determination by consumers. Section 2.2 presents the focus areas and the case-studies of this dissertation. Section 2.3 gives a short overview of the research approach.

2.1. Evidence production for food safety

2.1.1. Risk analysis as regulatory framework for food safety

In this section, the development of risk analysis principle is elaborated because it is the central tool to determine relevant evidence in food safety in the EU (EC, 2020; EC & EP, 2002).

2.1.1.1. The risk analysis concept

The risk analysis concept was defined and introduced in the 1980s after criticism on technocratic approaches in the regulation of risk by the National Research Council in the US. Beginning from the US, the risk analysis model established as state-of-the-art practice in food safety determination (Millstone, 2009). It consists of three components, risk assessment, risk management, and risk communication as displayed in Figure 2.

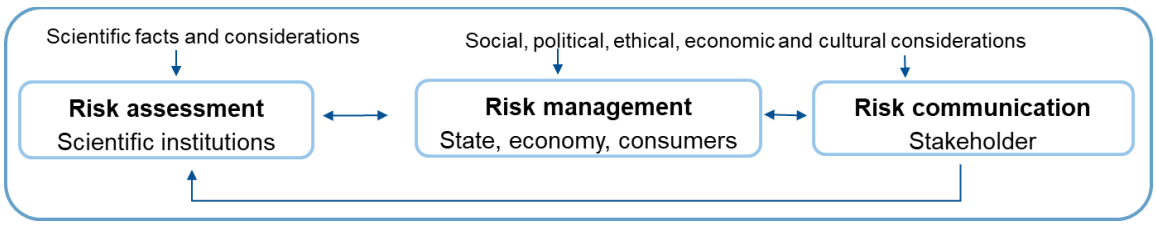


Figure 2: The practice of risk analysis (Source: Millstone, 2009)

The National Research Council describes risk assessment as “the use of the factual base to define the health effects of exposure of individuals or populations to hazardous

materials and situations [which should be] independent from political, economic, technical considerations that influence the design and choice of regulatory strategies" (NRC, 1983). Risk assessments should be based on the best available scientific data. They should further be independent, objective, and transparent (CAC/GL, 2007; EC, 2020; Millstone, 2009). The process of a risk assessment includes four standardized steps, hazard identification, dose-response assessment (hazard characterization), exposure assessment, and risk characterization (NRC, 1983). The application and form within these steps vary by hazard[4]. Risk management is the second step in the risk analysis process. It should consider risk assessment and include additional considerations of the food chain, as well as social, economic, and political concerns. Additional information can include for example effectiveness of measures, feasibility, or environmental impact. Risk management includes the weighting of advantages and disadvantages. It should be transparent, consistent, and fully documented (CAC/GL, 2007; EC, 2020; Millstone, 2009; NRC, 1983). Risk management entails political and value-based decisions which is different to risk assessment. Risk management needs to determine acceptable risk levels that a society is willing to tolerate (Cunningham, 2005). The third element of risk analysis is risk communication. Risk communication is a tool to provide reliable information to interested parties and to increase the public understanding of complex risk issues. It aims to increase public trust, foster understanding of risk management decisions, enable stakeholder involvement, and improve the efficiency of risk analysis (CAC/GL, 2007; EFSA, 2018).

Risk analysis rapidly evolved as state-of-the-art practice in food safety determination as soon as it was included in the Codex Alimentarius[5], the WHO, the WTO, and the Organization for Economic Cooperation and Development (OECD) (Millstone, 2009). In the EU, the risk analysis model was implemented within the reorganization of the food safety system after the BSE crisis in the 1990s. It replaced an inconsistent, pragmatic, and in-coherent concept which was originally invented to harmonize EU markets (van der

[4] A detailed description of differences between case-studies is provided in section 2.2.5..

[5] The Codex Alimentarius Commission is a joint body of the Food and Agricultural Organization (FAO) and the WHO which aims to protect health and enable fair trade. It establishes the Codex Alimentarius, a collection of internationally adopted standards and practices for food safety. The Codex is included as reference standard in the SPS agreement (FAO/WHO, 2020).

Meulen & van der Velde, 2004; Vos, 2000). The implementation of the risk analysis principle was accompanied by institutional and procedural changes – most importantly the establishment of the EFSA as independent risk assessment authority and the separation of science and decision making (Szajkowska, 2012). The EFSA is responsible for independent scientific consulting, scientific reports, and recommendations for the European Commission, the European Parliament, and member states. Member states have own risk assessment authorities with different responsibilities (Henning et al., 2014). The complex risk analysis landscape is illustrated in Figure 3 using the example of PPP regulation. Different European and national bodies (in this example: German Federal Institute for Risk Assessment (BfR) and Federal Office of Consumer Protection and Food Safety (BVL)) are responsible for the approval of active ingredients, PPP formulations, and their MRLs. In total, there exist three complex and different institutionalized evidence practices in the EU for food safety regarding PPPs (excluding re-approvals).

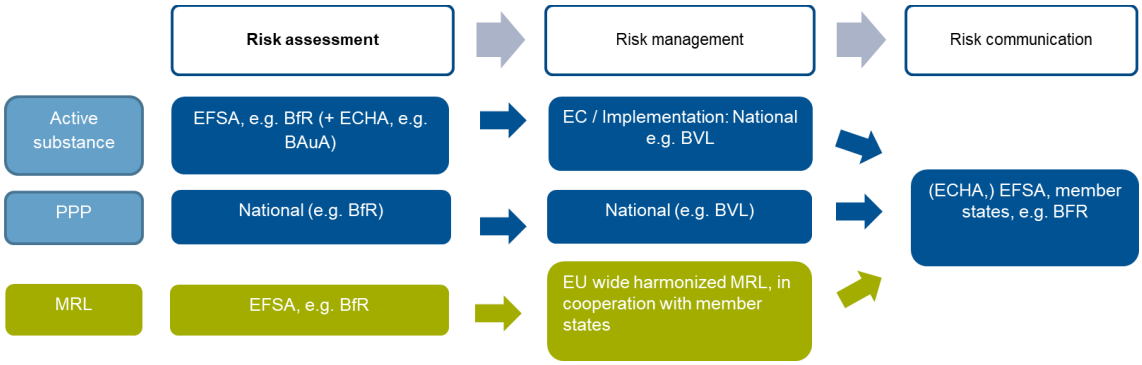


Figure 3: Responsibilities in EU food safety determination at the example of PPPs

2.1.1.2. Difficulties in risk analysis

The previous section shows the complexity but also the high relevance of the evidence practices of risk analysis in food safety determination. This complexity and relevance entails issues of different nature. A detailed analysis of specific issues in the case of PPPs in provided in study 1. Thus, the following section gives a brief overview to exemplify some practical and conceptual issues of risk analysis. It will prove that the evidence practices and thus the relevant evidence for food safety is a controversial topic worth to investigate.

Practical difficulties in the scientific risk assessment originate in the availability and selection of data as foundation of evidence. It depends on the amount of available

information if a certain quality of data is accepted or not. Further, data can be contradictory and need to be evaluated and weighted (Dreyer & Renn, 2009; Hassauer & Roosen, 2020; Rhomberg et al., 2013).

Risk assessment needs to fulfill various high expectations: It should ideally be an objective provider of value-free information but objective scientific facts do not exist (Vos, 2000). However, in risk assessment, different value-laden decisions need to be made which do not only include scientific values but also extra-scientific values like in cases of uncertainty in risk assessment (Vareman & Persson, 2010). Especially constructivist literature argues how risk assessments' framing, practices, and decisions are influenced by culture, society, or institutions (Klinke & Renn, 2002; van Zwanenberg & Millstone, 2000). This challenges the basic assumption in risk analysis and the societal expectation that science-based risk assessment and political risk management need to be independent (EC & EP, 2002; Vareman & Persson, 2010). The transition of evidence from risk assessment to risk management is difficult: The ideal situation is that science is policy-neutral but policy-relevant. Further, it is necessary that regulators understand scientific processes and uncertainties to use evidence most efficiently in regulations. There is always a certain amount of interpretation of scientific statements, instrumentalization of scientific uncertainty for certain arguments, and the existence of controversial scientific opinions to one topic (Jasanoff, 1990; Wagner, 2016; Zwietering, 2015). A strict separation of risk assessment and risk management is practically impossible, an efficient exchange is necessary (Ruzante et al., 2010; Vos, 2000). On the risk management level, it is not specifically defined how regulators should deal with scientific facts and with related uncertainty, how they should define acceptable risk, or the role of precaution (Vos, 2000). For example, in the SPS agreement, there is a missing definition of the amount of science for decisions and thus for international trade regulations. SPS Agreement Article 2.2 says that "Members shall ensure that any sanitary or phytosanitary measure (…) is based on scientific principles and is not maintained without sufficient scientific evidence (…)". The definition of sufficient is missing (Cunningham, 2005).

There is an additional fundamental conflict between scientific risk assessment and political risk management. A scientific investigation must always be a revisable, open-ended inquiry. Tools and techniques can change with scientific progress. Methods depend

on the type of science. Methods and hypothesis have to be accepted by the scientific community which is itself a "value laden judgement". It is a challenge that science is evolving, while regulations generate fixed standards (Wagner, 2016). This problem is amplified because the development of new food technologies can be faster than scientific observations which makes it hard to develop suitable regulations for new hazards. Additionally, regulations are expensive and work intensive. Once they are implemented, they will not change easily (IoM & NRC, 2003). In risk management, different mechanisms influence the setting of optimal regulations: A diffuse consumer public and a highly technical scientific inquiry (Trumbull, 2006). Even if science delivers optimal solutions for hazards, they could be too expensive or unsuitable. Optimum levels for societies are often stipulated by socio-economic issues, i.e. not only by scientific possibilities but also the willingness to pay. Everyone is involved in buying food and interested in low prices. Usually, only few people are harmed by food hazards under current food safety regulations. It is preferred to implement measures which have no significant impact on food prices (IoM & NRC, 2003).

The preceding criticisms have been raised since the invention of risk analysis in European food safety determination (Morvillo, 2020; Vos, 2000). Lately, they became publicly relevant during the debate on the re-approval of glyphosate in Europe in the 2010s. The debate emerged from scientific controversy, the dominant role of expert knowledge and related criticism, and a lack of trust in the regulatory system (Morvillo, 2020). These criticisms show that the evidence practice(s) for food safety in the EU are contested societal processes with outcomes which cannot be purely objective. They come with practical difficulties, different conceptual interpretations, ambiguities, and the question when scientific output is sufficient to become evidence in the sense of societally accepted and relevant knowledge for safety. This thesis aims to investigate these facets by analyzing evidence practices and provide insights in the construction of food safety.

2.1.2. Consumers and intuitive toxicology

Consumers constitute the second central part of the actors' network. Consumers are a central element in the determination of food safety and acceptability of risks from food but face fundamental difficulties in evaluation. In the last century, traditional knowledge was increasingly replaced with scientific knowledge (Spiekermann, 2011). Technical

advancements in food production make it difficult for consumers to use sensorial abilities to evaluate food safety regarding chemical contaminants or processing techniques. The result seems to be an increasing technology aversion (Lusk et al., 2014). Additionally, food markets are characterized by the close connection to health and differences between practical and "objective" knowledge (Spiekermann, 2011).

Scientists from different research backgrounds have analyzed the consumer evidence practice and used different approached to describe and research consumers food safety determination – for example intuitive toxicology (Kraus et al., 1992), risk perception (Slovic, 1987), or subjective food safety (Grunert, 2005). A detailed analysis and presentation of the knowledge on consumer evidence practice is provided in study 1. The following section gives a brief overview and shows how the recognition of the role of public risk perception changed in science as well as in policy.

The unrestricted credibility of expert risk evaluation started to get increasingly questioned after publications like "Silent spring" by Rachel Carson in the 1960s[6]; additionally, different environmental disasters such as Tschernobyl or Bophal showed that technological progress is likely to produce unintended negative effects. Sociologists like Ulrich Beck started to describe the definition of risks and their acceptability as depending on social definitions and perceptions (Beck, 1986). In parallel, scholars began to acknowledge and analyze consumers' different ways of risk evaluation. For example, Starr drew attention on the topic of public acceptability of risks. He recognized that acceptability is not only a risk issue but also a question of related benefits and voluntariness (Starr, 1969). Additionally, research started to address the question why lay people come to different conclusions about risks than experts. The term "intuitive toxicology" for chemical hazards exemplifies how literature started to recognize that consumers have their own way to evaluate risks which goes beyond likelihood and probability (Kraus et al., 1992). In his pioneering work, Slovic was able to proof that risk perception is multidimensional and that consumers evaluate hazards based on characteristics like *dread* and *familiarity* (Slovic, 1987). These dimensions were confirmed for a variety of hazards including food hazards like pesticides or

[6] The book critically discusses the use of pesticides and their impact on ecosystems and health (Carson, 1962).

microbiological hazards (Fife-Schaw & Rowe, 1996; Sparks & Shepherd, 1994). In the last decades, consumer risk research shifted from the origin of misunderstandings and "misperception" to the correction of the subsequent knowledge deficits and risk communication research. These studies focus on the difference in risk evaluation by expert and lay people and related correction of these differences (Hansen et al., 2003; Kraus et al., 1992; Lofstedt, 2006). A major objective was to understand and avoid the social amplification of risks (Kasperson et al., 1988; Lofstedt, 2006). Related is research on trust as it was identified as a critical element in the effectiveness of risk communication (Lofstedt, 2006). Analogous, the first discussions and analyses of individual factors in risk perception emerged, for example on cultural differences in risk perception, risk as feeling, or the role of heuristics and biases in perception (Douglas & Wildavsky, 1983; Loewenstein et al., 2001; Lusk et al., 2014; Sunstein, 2002). In recent years, this stream of literature gained in importance. There has been an increasing amount of literature which focusses on the consumer as individual who is influenced by worldviews, moral judgements, emotions, individual characteristics, or identity definition (Haidt, 2012; Hansen et al., 2003; Kahan, 2016a; Lusk et al., 2014). Over the last five decades, there has been a shift in research from the role of hazard-linked characteristics to the role of perceiver-linked characteristics in risk perception as well as a shift from the analysis of expert-lay gap to the underlying psychological and sociocultural mechanisms in decision making (Bouyer et al., 2001; Lusk et al., 2014). Nevertheless, thus far, nearly all studies rely on the assumption that there exists an objective, empirically derived level of safety which is perceived by consumers differently for various reasons – varying between hazards or individuals. That is exemplified in terms such as a "perception" or "amplification" of risk.

This assumption has important implications for the role of consumers in the evidence practice of food safety. The term "amplification", for example, implies an overestimation of risks compared to "objective" risk assessments. In general, research on consumer perception and intuitive toxicology-initiated discourses about the recognition and role of risk perception by consumers. This is closely related to the discussion about the sovereignty over the general definition and the acceptability of risk in societies which lasts until today.

Scholars describe the implications of their research mostly on the level of risk communication to avoid amplification or misperception (Fife-Schaw & Rowe, 1996; Kahan et al., 2007; Kaptan et al., 2017; Lofstedt, 2006; Lusk et al., 2014; Sunstein, 2002), the inclusion of the public in risk management and communication (Klinke & Renn, 2002; Lofstedt, 2006; Slovic, 1987), to forecast public perception for upcoming hazards (Slovic, 1987), to identify and consider biases in decision making (Sunstein, 2002), to recognize different worldviews in risk regulation (Kahan et al., 2007), to explain differences between individuals with certain characteristics (Bieberstein & Roosen, 2015), or to "mitigate real risks and irrational fears" (Loewenstein et al., 2001).

The recognition of the relevance of consumer perception is also visible in the changing role of consumers in EU policy. For a long time, consumers were only seen as victims who need to be protected, also through the scientific evaluation of food hazards (Covello & Mumpower, 1985). Currently, the EU food law treats consumers as "reasonably intelligent, responsible and capable of making informed choices" (van der Meulen & van der Velde, 2004). Consumer protection includes the opportunity for consumers to make informed choices and highlights risk communication as a central element in risk analysis (van der Meulen & van der Velde, 2004). Nevertheless, these perspectives limit consumers as a reactive element in the process of food safety determination. Risk perception is not necessarily seen as valid evidence for safety. Based on the trading zone concept, this dissertation assumes the consumer practice as interacting but independent from scientific food safety determination.

2.2. Focus areas of the dissertation

This dissertation focusses on different evidence practices in food safety based on the trading zone concept. Four studies provide different perspectives on the cases which are described in section 2.2.1. to 2.2. 4.. Study 1 analyzes theoretical and empirical knowledge on food safety criteria in the context of evidence practices and provides contributions to a multidisciplinary conceptualization of the evidence determination process of food safety. Study 2 analyzes food safety criteria used by consumers. Study 3 analyzes the practice of risk assessment and its interaction with stakeholders. Study 4 analyzes the role of evidence practices as decision tools in risk management. The dissertation provides insights in the practice of evidence determination for food safety.

Two main case-studies are used, PPPs and microbiological hazards. In section 2.2.5., the reasons for selection and the differences between the cases are described.

2.2.1. Study 1 – Conceptualizing food safety – a multidisciplinary perspective on food safety criteria

Study 1 takes a conceptual, holistic perspective on the two different evidence practices. Food safety determination is a constant process of discourse, even in areas with a long history of risk analysis like for PPPs. That is exemplified in the recent debate on the re-approval of Glyphosate in Europe (Morvillo, 2020). So far, literature explains this phenomenon by framing effects, misperception, or amplification mechanisms from different societal actors but usually assume that it is possible to determine one value-free safety level empirically. That assumption was proven wrong for example in the scientific debate on the assessment of the carcinogenicity of Glyphosate which was not consistent across different scientific organizations. Contrasting to assumptions of previous literature, the trading zone concept based on different evidence practices indicates that there is not a clear answer on how to define food safety. It rather points towards the existence of a variety of different criteria, which are used by relevant actors to determine "safe" food. Study 1 focusses on that perspective. It aims to provide a conceptualization of food safety determination based on the analysis of the two evidence practices, science-based risk analysis and consumer evaluation.

For food safety determination, an extensive base of literature exists from different research fields – for example, toxicology, risk research, psychology, or economics. Different research fields have different understandings of the term safety. From these fields, food safety criteria are collected and integrated into a conceptual framework.

The specific contributions of study 1 to the dissertation are listed in Table 1 on the next page. The study provides the conceptual basis for this dissertation, including a detailed analysis of different research fields and evidence practices.

Conceptual frameworks are a valuable tool in risk research. They are a relevant instrument in the analysis of "concepts, principles, theories, frameworks, approaches, methods, and models" (Aven, 2018). Additionally, conceptual frameworks have a close link to applied research and can provide framing or interpretations for risk phenomena,

hypothesis for empirical risk research, and approaches to optimize risk analysis (Aven, 2018).

Table 1: Research questions and contributions of study 1

Research question	Contribution of Study 1, sub-questions
A) How does the determination of sufficient evidence for food safety differs between the science based risk analysis process and intuition-driven consumers?	• Collection of food safety criteria from both sides to describe "food safety criteria" within a framework instead of creating a universal definition. • Which criteria are used by risk analysis and consumers in their evidence practices to determine safe food and which difficulties are attached?
B) How do these evidence practices interact in the determination of food safety and influence the selection of evidence for food safety?	• Collecting food safety criteria from both sides. • Identifying similarities, zones of contact, challenges.

2.2.2. Study 2 – Food safety criteria used by consumers – a qualitative analysis of in-depth interviews

Study 2 focusses on the understanding of food safety criteria used by consumers.

Based on the trading zone concept it is assumed that consumer food safety determination is based on an independent evidence practice including own categories and meanings of food safety criteria. It is an important task to identify which food safety criteria are relevant for consumers for an efficient inclusion of consumer evidence in risk analysis. Contrasting to food safety criteria from natural sciences, consumer food safety criteria are less tangible and a product of complex societal and psychological processes. So far, literature identified multidimensional aspects and various factors in consumers' food safety perception and determination processes. Nevertheless, many studies are based on existing concepts, use unsuitable measuring instruments, or focus on very specific aspects in these processes. Study 2 takes a step back and employs an inductive approach to collect and classify food safety criteria used by consumers beyond existing concepts. That takes account of the assumed independent evidence practice and allows a holistic perspective on consumer food safety criteria. That might function as a basis for optimized understanding and integration of consumer food safety determination in risk analysis.

The specific contributions of study 2 to the dissertation are listed in Table 2. The study provides insights in the consumer evidence practice.

Table 2: Research questions and contributions of study 2

Research question	**Contribution of Study 2, sub-questions**
A) How does the determination of sufficient evidence for food safety differs between the science-based risk analysis process and intuition-driven consumers?	• Identify and classify the criteria that consumers use to assess and determine food safety. • Holistic view on the consumer evidence practice.
B) How do these evidence practices interact in the determination of food safety and influence the selection of evidence for food safety?	• The understanding of food safety criteria optimizes evidence-based inclusion of public food-safety determination in policy through adapted measurements, problem structuring, and framing of food-safety assessments.

2.2.3. Study 3 – Practice of food safety determination – The integration of stakeholder evidence in risk assessment

Study 3 takes a perspective on the evidence practice of risk assessment and its contact zones with consumers and food production.

Based on the trading zone concept, it is assumed that the conceptual differences in evidence practices have an impact on food safety determination practice. Study 3 focusses on the regulatory practice of food safety determination, more specifically on risk assessment in the EU. It analyzes two cases, PPPs and microbiological hazards. Risk assessment is a key element in European regulations and "should be undertaken in an independent, objective and transparent manner" (EC & EP, 2002). It often is assumed to be an ideal representation of scientific values. That may produce unrealistic expectations, as well conceptual and practical conflicts (Vareman & Persson, 2010).

Risk assessments need to interact with stakeholders in certain ways. The amendments of the General Food Law in Regulation 2019/1381 aim for an early interaction with stakeholders in risk assessment (Chatzopoulou et al., 2020). These stakeholders may have different understandings or types of evidence for food safety. The study aims to investigate the contact zone between science and stakeholders and thus conflicts or co-

production between evidence practices. It is of interest, how risk assessors integrate stakeholder evidence in their work and which conflicts arise from that.

The specific contributions from study 3 are presented in Table 3. The study provides insights into the practice of food safety determination.

Table 3: Research questions and contributions of study 3

Research question	**Contribution of Study 3, sub-questions**
A) How does the determination of sufficient evidence for food safety differs between the science-based risk analysis process and intuition-driven consumers?	• Relevant differences in the practice of risk assessment. • How do differences between evidence practices influence, complement, and challenge the practice of risk assessment? • How do risk assessors deal with these differences in practice?
B) How do these evidence practices interact in the determination of food safety and influence the selection of evidence for food safety?	• Interaction dynamics of risk assessment organizations stakeholders. • How do risk assessors interact with stakeholders in selecting and using evidence and thus contribute to the in- or exclusion of (societal) information?

2.2.4. Study 4 – Cost-benefit analysis as decision tool in evidence-based policy

Study 4 focusses on the use of evidence in risk management, more specifically the cost-benefit analysis, and describes it as an evidence practice itself.

Based on the trading zone concept, it is assumed that evidence for food safety regulations is negotiated and codified in a trading zone. Study 4 focusses on the practice of these negotiation processes in risk management. Regulation in health- and environmental policy are products of complex processes which include the consideration of different risks, groups, interests, and target conflicts. In risk management, there exist different procedures to evaluate and weight different forms of evidence, for example the cost-benefit analysis.

This literature-based study aims to describe evaluation and weighting mechanisms in cost-benefit analysis, its establishment in decision processes as evidence practice, and critically discuss its problematic aspects.

The specific contributions from study 4 are presented in Table 4. The study provides insights into the negotiation processes of different forms of evidence.

Table 4: Research questions and contributions of study 4

Research question	**Contribution of Study 4, sub-questions**
A) How does the determination of sufficient evidence for food safety differs between the science-based risk analysis process and intuition-driven consumers?	• How do differences between evidence practices influence, complement, and challenge the practice of risk management? • How do risk managers deal with these differences in practice?
B) How do these evidence practices interact in the determination of food safety and influence the selection of evidence for food safety?	• How do risk managers negotiate different forms of evidence in cost-benefit analysis?

Focusing on the areas described in the previous sections, this dissertation aims to investigate when and how knowledge for safety is considered as evident (practicing evidence), and furthermore which evidence is transferred into the risk analysis process (evidencing practice).

2.2.5. Case studies

The field of food safety determination covers a broad range of hazards with different characteristics. Because it is not possible to cover all these processes, for this thesis, two case-studies are selected as main focus areas. The selection of case-studies of contrasting nature allows for extensive qualitative observations. Selected cases are (1) PPPs and (2) microbiological hazards. Study 1 uses the case of PPPs. Study 2 uses both hazards and irradiation as additional case. Study 3 uses both hazards, and study 4 employs some examples from the regulation of PPPs. The risk analysis concepts for both cases are based on same basic structure but differ on the levels of hazard characteristics and thus in scientific risk assessment practice, regulation, and consumer perception. The differences between case-studies allow a broader view and the identification of differences and similarities.

2.2.5.1. Hazard characteristics

The case-studies differ on the level of hazard characteristics which are the basis for differences in risk analysis practice and risk perception.

A main difference is that PPPs, in contrast to microbiological hazards, are intentionally produced and introduced into the food chain. They are required to ensure food quality and production but produce risks in form of unintentional risks in the food chain and the environment. PPPs are hazards with multiple effects relevant on different policy levels: local effects for example on biodiversity, national effects for example on groundwater, or global effects for example on food safety (IoM & NRC, 2003; Sexton et al., 2007). Additionally, PPPs produce "transgressive effects on other systems outside the system of origin" (Renn et al., 2020), for example PPPS may accumulate in the food chain in form of PPP residues or metabolites in fish (Verbeke et al., 2005). Along with others, these characteristics are indicators for the classification of PPPs as systemic risk. Systemic risks are challenging to address in risk management and require interdisciplinary approaches and a careful determination of socially acceptable safety levels (Renn et al., 2020). PPPs have specific exposure characteristics: They have a constant potency or degrade, in mixtures they are reduced by dilution but can cause chemical reactions. Therefore, it is most important to control the raw material (IoM & NRC, 2003).

Microbiological hazards are naturally occurring and are unintentionally present in the food chain. The number of pathogens can increase or decrease, depending on conditions in every step of the food chain, from primary production to consumption. In mixtures, pathogens can spread, therefore it is important to control the whole food chain (IoM & NRC, 2003). This is even more difficult because same foods products are produced from different companies, have different transport or storage conditions, and preparation depends on the consumer (van Schothorst, 2002). Raw products as well as processed food are traded within and across borders (Manfreda & De Cesare, 2014). Some pathogens are only hazardous for consumers in certain strains (for example certain strains of E.coli), in certain food products, or only for vulnerable groups, for example young children (van Schothorst, 2002). Besides acute effects, some pathogens can cause severe long-term effects. For example, in humans, the pathogen Campylobacter often causes acute effects like Gastroenteritis but also long term sequelae (Ruzante et al., 2010).

2.2.5.2. Scientific and regulatory practice

In food safety, "there is a significant divergence concerning risk assessment processes" (Chatzopoulou et al., 2020). These differences in risk assessment processes between the two cases are described in the following section. They are listed in Table 5, broken down by the four steps of risk assessment: Hazard identification, dose-response assessment, exposure assessment, and risk characterization. Table 5 illustrates the effect of hazard characteristics. For example, for PPPs, a risk assessment is initiated through intentional introduction and in contrast in the case of microbiological hazards through public health concerns (IoM & NRC, 2003). Exposure assessment for microbiological hazards is often based on complex modelling due to complex exposure scenarios, thus a standardization is difficult (van Schothorst, 2002). Additionally, the case-studies differ in used types of data, outcomes, and objectives. For example, for microbiological hazards, it is intended to estimate the risk for the population to develop appropriate control strategies. For PPPs it is intended to determine safety levels and to reduce exposure (IoM & NRC, 2003).

Table 5: Differences in risk assessment of Microbiological hazards and Chemical hazards (EC & EP, 2009; IoM & NRC, 2003; van Schothorst, 2002; Zwietering, 2015)

	Microbiological hazards	**PPPs**
Hazard identification	• Initiated through public health concerns, outbreak • Uses epidemiological data • Aim: Identification of organism	• Initiated by approval process • Uses in-vitro systems, animal studies, genomic sciences • Aim: Approval of substance
Dose-response assessment	• Outbreak data, usability of animal studies limited because of host-specificity • No outcome like NOAEL because single cell may produce illness • No clear link to biological mechanisms • Response of population more variable	• Animal studies according to Good Laboratory Practice • Outcome: Dose considered as safe (NOAEL/NOEL[7]) • Clear link to biological mechanism • Response of population less variable

[7] No Observed Effect Level (NOEL) and No Observed Adverse Effect Level (NOAEL) are animal data based thresholds which are extrapolated to humans in form of the Acceptable Daily Intakes (ADI) and Acute Reference Doses (ARfD) (Benford, 2000).

	• Expert judgements necessary	
	• Various endpoints can be used • Data from multiple studies can be combined • High degree of uncertainty	
Exposure assessment	• Modelling movement in time and space • Frequency and number of ingested microorganisms • Expert judgement necessary, because industry data often not available	• Determine the fracture with is absorbed and bioavailable in body
	• Contribute different ways of routes: inhalation, dermal, oral • Adverse effects on humans resulting of one substance	
Risk characterization	• Harmfulness of pathogen • Risk for population from certain food • Probability distributions of the variability in illness • Risk rankings	• Determining the dose which is not harmful to human, environment, etc.

In the EU, introduction of PPPs and presence of residues are regulated separately to address the approval and the occurrence of potential harmful effects. Both risk assessment processes are highly standardized and require the provision of specific data sets produced after Good Laboratory Practice (EC & EP, 2005b, 2009). EFSA assesses applications for new products in specific working groups. Based on these assessments, a standing committee in the European Commission decides about approval (Chatzopoulou et al., 2020).

For microbiological hazards, risk management often needs to act *ad hoc*. Risk managers need to set priorities for (newly occurring) hazards which should be analyzed in risk assessment to establish measures to protect consumers. Priority setting needs to include societal factors. Risk managers use the results of risk assessments to decide about or to define actions (van Schothorst, 2002). This includes a weighting of different factors like health protection, ethical or legal levels, or free trade (Manfreda & De Cesare, 2014; Zwietering, 2015). These actions can include the definition of specific food safety objectives or food safety criteria as threshold levels for pathogens in certain food products at the time of consumption (Membré & Guillou, 2016; van Schothorst, 2002). An important management instrument is the hazard analysis and critical control points (HACCP) principle. Since the 1990s, it is a standard management tool in microbiological

food safety (Zwietering, 2015). It aims to prevent food products from unacceptable contamination instead of inspecting the product (Manfreda & De Cesare, 2014; van Schothorst, 2002).

In EU regulations, food safety criteria are defined for specific products, for example thresholds for *Salmonella* concentration in minced meat. Additionally, information on analytic reference methods (usually ISO-standards) and potential actions for high pathogen concentrations are defined. Generally, food producers are responsible to comply with the food safety criteria but are able to apply individual actions based on HACCP principles and good practice, for example the Good Hygiene Practice. Important is the correct documentation and labelling to ensure traceability (EC & EP, 2005a; Manfreda & De Cesare, 2014). It is criticized, that the definition of strict threshold levels in EU regulations does not comply with the high variability and uncertainty in current quantitative microbiological risk assessment approaches (Zwietering, 2015). Additionally, food safety criteria in EU regulations do not comply with Codex Alimentarius Commission guidelines which require the determination of risk-based food safety objectives and performance objectives (Manfreda & De Cesare, 2014).

2.2.5.3. Risk perception

The case-studies cover two types of food hazards differentiated by consumers: Technical (more likely to be overestimated) and naturally occurring (more likely to be underestimated) hazards, or more specifically chemical and microbiological hazards (Kaptan et al., 2017; Yeung & Morris, 2001). Literature provides different explanations for such differences in risk perceptions, for example Slovic (1987) introduced the concept of *dread* and *unknown* as factors to express risk perception, including control, fatality of consequences, effect on future generations, risk equality, risk development (dread) as well as observability, long-term effects, knowledge of exposed people and experts (unknown). In different studies, PPPs score high on the *dread* dimension and medium to high on the *unknown* dimension. Microbiological hazards score low on the *unknown* dimension but high on the *dread* dimension (Slovic, 1987; Sparks & Shepherd, 1994). Additionally, risk perception depends on benefit considerations or moral concerns which can differ between cases (Kaptan et al., 2017; Starr, 1969). As explained in the previous sections, it was shown that hazard characteristics are only one factor in risk perception besides individual

characteristics (Siegrist, Keller, et al., 2005). A more detailed analysis of these factors is provided in study 1.

The previous section illustrates that the evidence practices in safety determination vary on many levels even if both cases rely on the same principle. The selection of the different cases allows for a broader perspective on evidence practices in food safety.

2.3. Research approach

This dissertation aims to analyze evidence practices in food safety. It proposes a new perspective on food safety determination and is of an explanatory and interpretative nature. A combination of different qualitative methods is used to address the research questions elaborated in the previous sections: Qualitative conceptual framework building, qualitative analysis of in-depth interviews, an expert interview study based on a case-study approach and a literature study. All methods are used as a triangulation to provide different perspectives on the research questions described in section 1. Triangulation does not aim to provide complementary reproductions of results but different aspects or constructions of phenomena (Flick, 2011). The qualitative perspective allows to account for the complexity of food safety determination in a societal context, a consideration of framework conditions in which it takes place, and the first application of the concept of evidence practices. A more detailed description of the specific research approaches including data collection and analysis is provided in the related studies in chapter 3.

3. Essays

3.1. Toward a conceptual framework for food safety criteria: analyzing evidence practices using the case of plant protection products

Publication: Hassauer, C., & Roosen, J. (2020). Toward a conceptual framework for food safety criteria: Analyzing evidence practices using the case of plant protection products. *Safety Science*, *127*, 104683. https://doi.org/10.1016/j.ssci.2020.104683

3.1.1. Abstract

Food safety has a significant influence on food markets and is of great societal importance because it protects human health and life. Most studies presume that the presence or absence of food safety can be objectively assessed based on data from natural sciences and might be further interpreted and perceived in different ways; however, there is no consensus on the definition of food safety. Disputes within the scientific community and increasing public discourse suggest that there is no generally accepted definition of what is “safe” or “unsafe”. This paper introduces a framework that describes food safety in a broader sense, using the example of plant protection products, by identifying different evidence practices through the classification of criteria from various research fields. Data were collected in an integrative literature review. Criteria for assessing food safety were classified and collected within a conceptual framework that acknowledged the multidisciplinarity of knowledge bases. The analysis highlights the questions that arise when determining these criteria. We conclude that obtaining a generalized definition of food safety is not possible. Instead, our results showed the determination of food safety by criteria at different levels: science-based criteria at knowledge and value levels that result in standards and consumer-based criteria at knowledge and value levels that result in behavior. A better understanding of food safety criteria helps to show deficits in the current risk analysis practice and points to solutions for more consistent regulations, leading to more stable market conditions and a stronger mutual understanding.

3.1.2. Introduction

The regulation of food safety receives a great deal of attention in societal and political discussions because it ensures the protection of human health. In addition, it is of high economic relevance because of its role in non-tariff trade restrictions and the potential for high losses in cases involving food scandals. It is quite challenging for regulators to determine food safety because different societal groups appear to have a different understanding of what is meant by safety, as indicated by different evidence practices[8] in food safety negotiations; therefore, public discourse in safety negotiation is vulnerable to mistrust, misunderstanding, and mutual degradation, especially in critical cases, such as it was observed in the re-approval process of glyphosate in the European Union (EU).

Since the 1990s, the Sanitary and Phytosanitary Agreement of the World Trade Organization has determined which evidence is accepted as a justification for non-tariff trade restrictions based on the risk analysis process published in the 1980s by the National Research Council (Millstone, 2009; NRC, 1983). This process is based on a scientific risk assessment that quantifies the risk to human health. For plant protection products (PPPs),[9] most empirical data stem from the field of toxicology (EC & EP, 2009), and food safety regulations control the intentional introduction of PPPs in the form of approvals as well as the unintentional presence by determining maximum residue levels in food. PPPs are an interesting case-study for several reasons. As they are human-made, their introduction, use, and safety can be regulated (IoM & NRC, 2003). In addition, their risk analysis process is highly complex in the European context and PPPs elicit a high degree of skepticism by consumers. The food safety-related literature presumes that the presence or absence of food safety can be assessed in an objective way that might be further perceived, assessed, or interpreted, depending on, for example, the risk characteristics or values and worldviews of specific societal groups (Hansen et al., 2003; Kahan, 2016a; Slovic, 1987). However, the following three main arguments question this presumption.

[8] Evidence practices are understood as procedures of evidence generation (practicing evidence) and processes of using and embedding evidence (evidencing practice).

[9] The term includes product formulations and related active substances. We avoid the term ‘pesticides’ because it includes biocides, which are regulated differently in the EU.

First, we find a lack of consensus on the definition of food safety. Although intensively used, neither the scientific literature in natural sciences and consumer research nor regulators consistently define the term. Some authors define food safety as the absence of hazardous substances (Herges et al., 2017), while others refer to a specific certainty or probability of an adverse effect or even the inverse of risk (Henson & Traill, 1993; OECD, 1993). Given these differing definitions, two types of assessments are necessary to determine food safety: (1) the presence or absence of a hazard that can be determined using a hazard-assessment approach, and (2) the probability and severity of an adverse effect that is established using a risk assessment approach. Both approaches are used in European food safety practice with respect to PPPs: hazard-based approaches are used, *inter alia*, in classifications of carcinogenicity, and risk-based approaches are applied to determine maximum residue levels (Barlow et al., 2015).

Second, discourse among scientists and within scientific organizations demonstrates a lack of agreement in determining whether a food is safe or not. The recent discussion among various scientific assessment authorities on the inconsistent classification of the carcinogenicity of glyphosate exemplifies this issue (Portier et al., 2016). Inconsistencies might be caused by how the authorities deal with contradicting evidence and scientific uncertainty, which appears at every stage of the risk assessment process and is difficult to quantify (Barlow et al., 2015; Rhomberg et al., 2013).

Third, the mobilization of scientific knowledge in risk assessments is being increasingly questioned and discussed by the general public, which indicates that consumers use different criteria to evaluate what is and is not safe food. The assessment of food safety in terms of PPP residues is problematic for consumers because most negative effects are chronic, that is, caused by exposure to low doses over a long period of time. Neither PPPs nor their effects are usually detectable immediately following product consumption (Nau et al., 2002; Shaw, 2005). EU citizens are very concerned about pesticide residues in different food products (EFSA, 2010, 2019b). Consumers in Germany consider PPP residues to be the most influential hazard pertaining to food quality and safety, although the actual risk from PPP residues in food is rated low by scientists (EFSA, 2016b). This raises the question of whether consumers are using alternative criteria to determine whether a food product is safe, rather than overestimating numerical risk estimates.

In contrast to presumptions presented in the literature, the above statements indicate that there is no clear and generally accepted, value-free definition of food safety. Thus far, there appears to be no holistic consideration of the complex determination process in society; therefore, the aim of our study was to introduce a framework that describes food safety in a broader sense by identifying different evidence practices by collecting and classifying criteria from various research fields. It was built based on a systematic literature review using a grounded theory approach. The framework was developed conceptually and does not aim to describe theoretical relationships among criteria. It categorizes and summarizes multidisciplinary approaches that aim to describe various aspects in determining food safety. By comparatively analyzing either the risk analysis or the consumer part, the framework highlights the difficulties involved in determining these criteria and provides a more holistic description of food safety than would a universal definition. This exercise generates valuable insights into the different players involved in food safety discussions and might increase a mutual understanding and acceptance of different opinions. While not being able to arrive at a common framework across risk analysis and the consumer approach, the framework sets a basis for a more efficient interaction among regulators and consumers, and potentially for a better understanding of food safety regulations. A major objective of our study was to include consumers' food safety determination process not merely as a "critical interface between scientific facts and personal opinions and values" (Ropeik, 2011) but as a self-reliant evidence practice. This unconventional view of consumer food safety determination potentially provides new approaches to integration into policy.

Our paper is structured as follows: section 2 includes the guiding theory behind data collection and the process of building the framework; section 3 presents the criteria extracted from the literature as well as the developed conceptual framework; section 4 discusses the results and the implications for current food safety practices and research; and section 5 provides concluding remarks.

3.1.3. Procedure

The preceding statements suggest that determining food safety within a society is highly complex. We developed a conceptual framework by which to describe food safety not with a definition but with criteria. To that end, we defined a criterion as "a principle or

standard by which something may be judged or decided" (Oxford University Press, 2017). Building on literature from various disciplines, a conceptual framework offers the possibility of developing a network of linked concepts in food safety. Jabareen (2009) describes a conceptual framework as an appropriate method by which to describe "complex phenomena linked to different bodies of knowledge"; therefore, rather than provide a theoretical explanation of relationships among variables, we provide an understanding of the complex negotiation process involved in food safety. The systematic approach to the framework and literature review is based on the guiding theory of Galison's trading zones (2010) that we introduce in section 3.1.3.1. To build a framework, Jabareen suggested eight steps, which we summarized in three superordinate working steps as follows: collecting the data (section 3.1.3.2), building the framework (section 3.1.3.3), and validating the results (section 3.1.3.4) (Jabareen, 2009).

3.1.3.1. Guiding theory

As mentioned, we based our analysis on the concept of trading zones as introduced by Galison (2010), who uses the concept to understand how knowledge and language can be combined when two disciplines interact, integrating their respective concepts and contingencies into a new body of knowledge and evidence. In his work, Galison examined the interactions and negotiation processes between two different fields of science: physics and chemistry. Despite its comparatively confined initial application, the concept of trading zones can be used to study food safety because Galison also suggested applying it to explain the interactions between two disparate societal groups (Galison, 2010); therefore, a central hypothesis of our study was that there are two zones that influence and are influenced by the evidence practices of food safety: scientific-based risk analysis and the consumer, both of which interact within the food system trading zone (for details, see Appendix A). Because the definition of evidence is controversial, we describe it here not only as being based on data but also as one based on a social phenomenon, or, more specifically, socially accepted knowledge (Cartwright, 2006; Kelly, 2016).

Galison assumes that the two spheres of reference are not hierarchical but overlapping; therefore, it is presumed that scientific risk analysis does not comprise the source that is reflected by the consumers' evaluation of food safety, but that both scientific and consumer food safety practices intersect at some point. This implies that consumers do

not merely interpret the outcomes of the risk analysis process but have a self-reliant evidence practice. Within the trading zone, the definition of food safety is negotiated. Knowledge and criteria, but not necessarily their function and meaning, are exchanged. For PPPs, this implies that the relevance of knowledge taken from the natural sciences need not necessarily have the same meaning for food safety as does consumer perception, and vice versa, in the risk analysis process. Another characteristic of a trading zone scenario is an imbalance of power (Galison, 2010), which is also a characteristic of food safety determination. There appears to be an imbalance in favor of risk analysis because consumers are viewed as mere perceivers of scientific facts influenced and misdirected by various factors (section 3.1.4.3); therefore, the development of our framework is based on these two perspectives—first, science-based risk analysis, and second, consumer-based risk assessment—in order to provide a new perspective from which to negotiate the determination of food safety.

3.1.3.2. Data collection

The data that helped create the conceptual framework comprise the literature from various scientific disciplines. Data collection included mapping the data sources as well as extensively reading and categorizing the selected data (Jabareen, 2009). This was done within a systematic, integrative literature review that followed the system of Torraco (2005), who suggests it as a possible basis for building a conceptual framework. As opposed to Jabareen's first step, Torraco suggests beginning the synthesis of a conceptual framework by structuring the topic according to a guiding theory (Torraco, 2005); therefore, the original method was adapted because it allows for a more systematic approach to dealing with the extensive amount of literature on risk analysis and risk perception in food safety. To collect the data, an adaptation of Galison's trading zones concept was used as an underlying theory. Based on this concept, there were two main areas of interest for the review: first, the risk analysis process, including the scientific, technical risk assessment that was translated into risk management and policy (risk communication, normally considered part of the risk analysis process, was excluded because it is does not determine but rather communicates, food safety criteria); second, the main area of interest was consumers' evidence practices. Using the key words in

Figure 4, we identified the first relevant sample of scientific information, followed by a snowball procedure based on the literature identified in the first round.

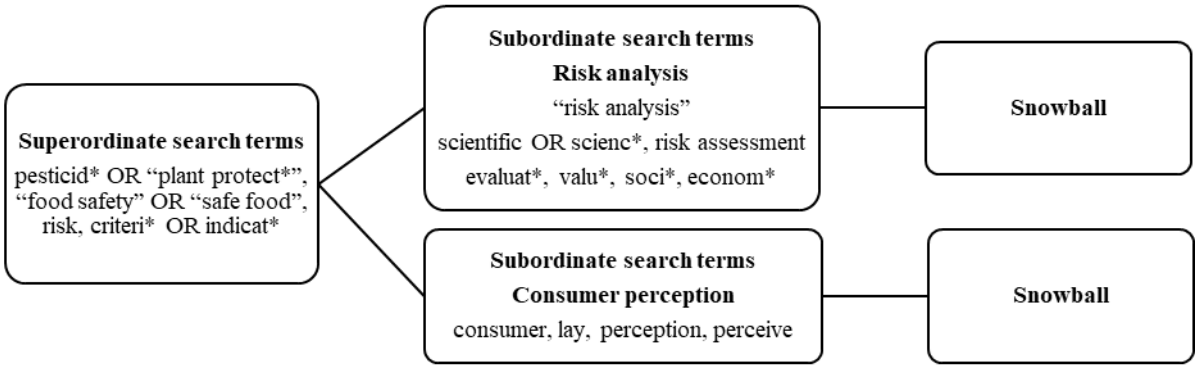


Figure 4: Search strategy of the literature review

The following exclusion criteria were used: not relevant, duplicate studies, related to outdated regulations, related to non-EU regulations, not related to PPPs (risk analysis context), environmental risk assessment, and conditions not comparable to EU (cultural, agricultural). The databases used were *ScienceDirect*, *Emerald*, and *Scopus* for scientific peer-reviewed articles or book chapters, and *Google* for gray literature from relevant organizations in risk assessment and management. The search was conducted in 2017 and 2018, and the data were organized in *Mendeley*. We identified 153 pieces of relevant scientific information, 78 of which were categorized as "risk-analysis"-related, 67 of which were categorized as "consumer"-related, and 8 of which were categorized as related to both concepts. In total, 120 of the identified sources were peer-reviewed papers; the other sources were books, regulatory documents, or gray literature.[10]

3.1.3.3. Building the framework

The steps for building the framework comprised identifying and naming the relevant existing concepts, deconstructing and categorizing the concepts, and integrating and resynthesizing the new concepts within the new framework (Jabareen, 2009). The concepts were categorized in two steps. The first step included a general categorization of the data on risk analysis and consumer-related evidence practices. The second step

[10] A detailed description of the distributions of publications by year and type is provided in Appendix B.

comprised building more detailed deductive and inductive categories. The final categories and levels of the framework, as well as the final classification logic, are described in 3.1.4.

3.1.3.4. Validation

The last steps in building the framework comprised validating and reevaluating the process (Jabareen, 2009). As a validation method, Jabareen suggests that results be presented at a conference of experts to enable them to add to or rework unnoticed aspects; thus, the results were duly presented at an agricultural economics seminar, a university research colloquium, and at the 27th Annual Conference of the Society for Risk Analysis – Europe in Östersund, Sweden, in June 2018. To obtain feedback from the perspectives below, the three presentation formats had different target audiences: agricultural and resource economists, consumer researchers and risk analysts. Feedback discussions comprised the following points:

- discussion of the role of international organizations in the framework,
- discussion of the naming of consumer criteria,
- discussion of the differentiation of consumer criteria, and
- comprehension questions.

Following feedback, the lowest level of consumer practice was changed from "assessment criteria" to "knowledge criteria". In addition, criteria definitions were refined because they were too unspecific. We identified some limitations in the suggested validation practice; these are discussed in section 3.1.5.

3.1.4. Results

During the review, we identified the criteria for food safety within the following two different evidence practices: risk analysis (section 3.1.4.2) and consumer evaluation (section 3.1.4.3). The criteria are classified into different levels that are linked within a conceptual framework (section 3.1.4.1). Both practices interact within the societal trading zone where food safety is negotiated. Because the focus of this framework is on identifying criteria and not on the negotiation process, these interactions are not closely addressed.

3.1.4.1. Conceptual framework for food safety criteria

The conceptual framework for food safety criteria comprises different criteria and related processes used to determine food safety (Figure 5). Here, we introduce the overall framework. All components are elaborated and explained in subsequent sections.

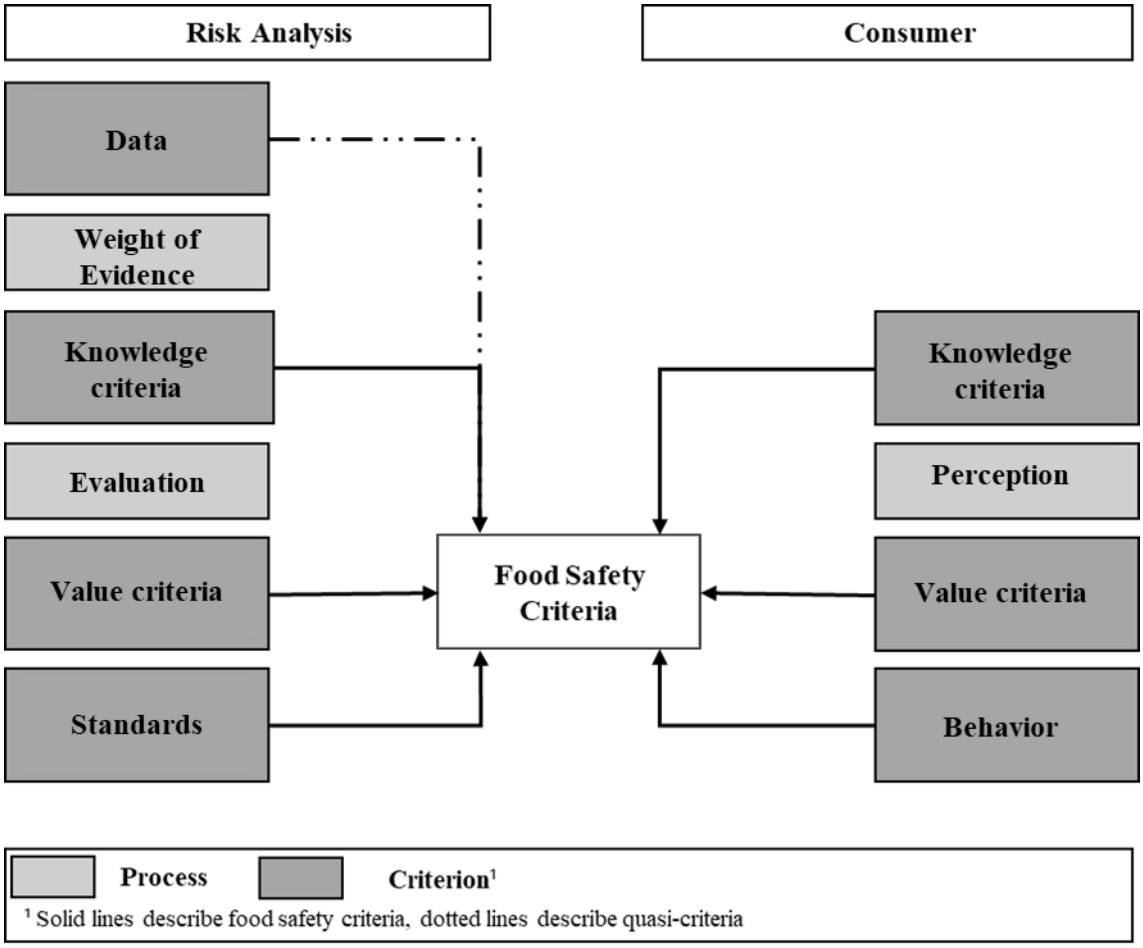


Figure 5: A conceptual framework for food safety criteria

Risk analysis (section 3.1.4.2) is based on data from natural sciences, which are problematic to classify as criteria under the applied definition. The Weight of Evidence process transforms data into knowledge criteria determined in institutionalized risk assessments. An evaluation process, either by risk managers or by socioeconomists, leads to a determination of safety in the form of value criteria. These are translated into legally binding standards comprising the final determination of safety in risk analysis.

Consumer safety determination is also based on knowledge criteria (section 3.1.4.3), which describe the relevant underlying dimensions of safety and are further perceived

and evaluated in the perception process. The result of this evaluation is the determination of a value criterion in the form of a stated concern or decision about whether something is safe. Lastly, safety is determined by the behavior of a consumer and his/her decision to purchase products perceived as containing PPP residues.

Both sides of the framework are characterized by similar levels of criteria: the knowledge and the value levels. Additional evaluation processes that result in stated value decisions about safety are observable and are fixed within regulatory actions or behavior. We now offer details about the different criteria.

3.1.4.2. Evidence practice of risk analysis

Expanding on the classification of Dreyer and Renn (2009), who divided the claims used in risk analysis into knowledge and value, we categorized the (quasi-) criteria of risk analysis within the levels presented in Figure 6 as follows: “data”, “knowledge criteria”, “value criteria”, and legally fixed “standards”. Between the levels, we present processes that transform one criterion into another. These include the Weight of Evidence approach and the evaluation, both of which are discussed below.

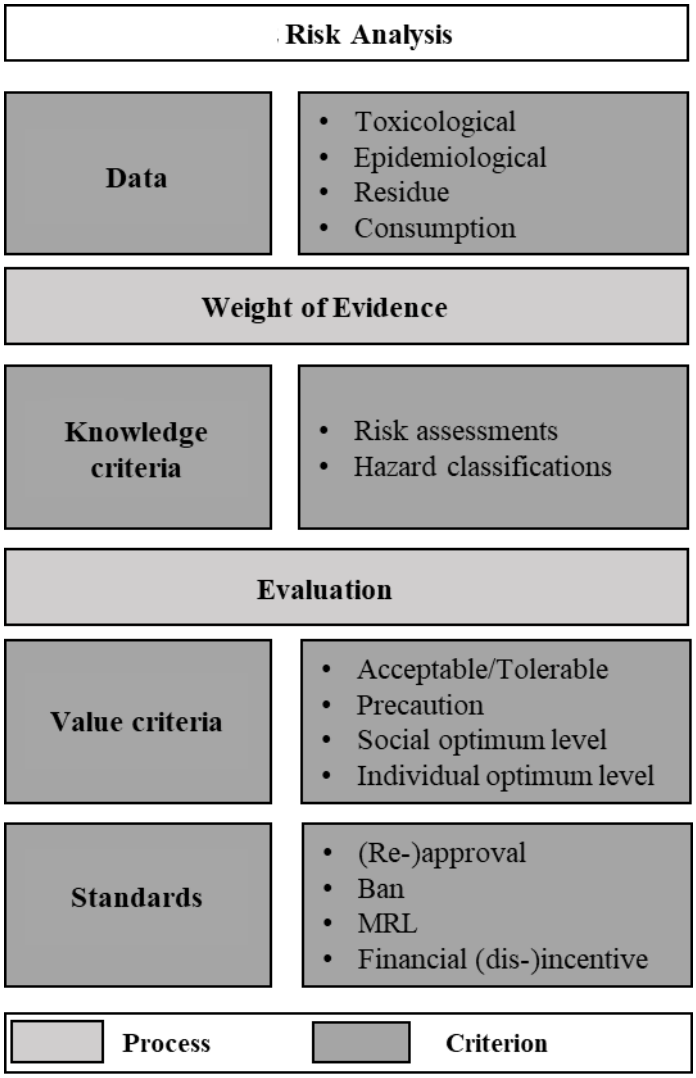


Figure 6: Food safety criteria in the evidence practice of risk analysis

Data

Data are defined as pieces of information (e.g., a single scientific study) (EFSA, 2015b). Required data in the evaluation of PPP safety include toxicological and/or epidemiological data to determine the effects and residue and consumption data for exposure assessments. During this stage, there is not necessarily a consensus on what is and is not safe. This makes it difficult to classify data as a criterion for food safety if such a criterion is defined as "a principle or standard by which something may be judged or decided" (Oxford University Press, 2017). Data might also be contradictory or insufficiently available. Various approaches exist to reflect the availability of knowledge (Douglas & Wildavsky, 1983; Dreyer & Renn, 2009). These include the assessment of available data based on the following parameters:

- Level of definition of outcome, level of definition of probability (Dreyer & Renn, 2009) and
- level of consent, level of knowledge about the future (Douglas & Wildavsky, 1983).

These classifications enable a critical view of the current risk analysis practice, a view that acknowledges that risk or hazard assessments are conducted within diverse states of knowledge that go beyond the required data in the approval dossiers for PPPs (EC & EP, 2005b, 2009). In contrast, technical risk assessment is viewed as only one way of evaluating safety that is applicable in all cases of well-defined probability and outcomes, consent, and appropriate knowledge about the future. In other situations, alternative solutions are necessary. For example, if there is consensus on an effect but the knowledge of its behavior in the future is uncertain, the problem is a lack of information and the solution is not a technical risk assessment but additional research (Douglas & Wildavsky, 1983; Dreyer & Renn, 2009). In risk assessments, data are often contradictory or not sufficiently available. In addition, we have observed that the amount of available data differs among cases; therefore, data relevance for the evidence depends on the ability to meet the objective of the assessment in question. Datasets might range from irrelevant to fully relevant depending on the available data. Irrelevant data might be used in assessments under consideration of uncertainty if more specific data are unavailable and hence become relevant (EFSA, 2015b). Incomplete data might stem from the existence of competing datasets, lack of perfect data, or the existence of competing theories (Rhomberg et al., 2013; Wagner, 2016). Data required for an ideal risk assessment are complex and, in most cases, not available (Reffstrup et al., 2010). Exemplary, issues of residue and consumption data are elaborated in Table 6.

Table 6: Issues in data

Kind of data	Issues and Literature
Residue data	• Variability due to agricultural practice[11] leads to uneven residue distributions in samples (Maclachlan & Hamilton, 2010; Tucker, 2008) • Overestimations of residues in supervised trials (Boobis et al., 2008) • Required field trials disregard heterogeneity of European agriculture and are only conducted in small numbers (Maclachlan & Hamilton, 2010) • Current PPP residue monitoring practice disregards cumulative and aggregate effects, levels of detection and quantification problematic for the estimation of cumulative effects[12] (Boobis et al., 2008) • Nonharmonized EU monitoring programs (Kennedy et al., 2015) • Average residue levels disregard inter-unit variability (Tucker, 2008) • Data are mostly available for agricultural commodities but not for processed food products (Boobis et al., 2008) • Different definitions for residues are used[13] (Boobis et al., 2008)
Consumption data	• Chronic intake assessments show methodological limitations through inadequate data[14] (Boobis et al., 2008; Tucker, 2008) • Nonharmonized collection of consumption data in the EU (Kennedy et al., 2015)

Weight of Evidence approach

As previously mentioned, it is likely that data are not sufficiently available and can be contradictory. To transform data into the next level of criteria, the Weight of Evidence approach comes into play. This approach is used in numerous contexts and by many institutions. Although the Weight of Evidence process is an essential component of the risk analysis process, we find inconsistencies in its definition and its systematic use within risk assessment institutions, which can lead to contradicting results even though they are based on the same amounts of available data. This is also the case for systematic reviews if they are used as synonyms for Weight of Evidence (Ågerstrand & Beronius, 2016; Haddaway & Bilotta, 2016; Pease & Gentry, 2016; Rhomberg et al., 2013; Weed, 2005; Whaley & Halsall, 2016). Although some institutions provide a clear characterization and description of their Weight of Evidence approaches (e.g., EFSA), others remain nonspecific and refer to the need for flexibility (e.g., ECHA) (Rhomberg et al., 2013).

[11] This includes different growing practices, spray equipment or growth stages of plants (Maclachlan & Hamilton, 2010).

[12] Residues under these levels might be problematic if cumulative or aggregate effects appear among different PPPs in a commodity and the non-reported PPPs sum to hazardous levels.

[13] Maximum residue level monitoring should be chemically simple and dietary risk assessment should include metabolites if they are toxicologically relevant (Boobis et al., 2008)

[14] Eating patterns can change over time and there is the issue of decreasing accuracy of recording over time (Boobis et al., 2008).

This can be problematic in PPP risk assessments when different authorities, here EFSA and ECHA, are involved (Barlow et al., 2015). In general, the Weight of Evidence approach relies strongly on expert judgment, especially in situations in which there are conflicting or insufficient data. This might lead to inconsistencies because experts can also be biased and must make value decisions in determining which data are accepted, which need to be collected, or how to deal with uncertainty (Fraiberg & Trebilcock, 1998; Nordlander et al., 2010; Van der Fels-Klerx et al., 2018; Wagner, 2016).

Knowledge criteria

The evidence practice of risk analysis is based on the criteria from natural sciences provided by assessment authorities. In the literature, we find the following different terms to determine these criteria: human health impact (Fantke et al., 2012), classic technical–scientific risk assessment (König et al., 2010), and knowledge claims (Dreyer & Renn, 2009). Knowledge criteria are considered to be useful for overcoming heuristic biases in policy making and for keeping the regulatory process in check (Fraiberg & Trebilcock, 1998; Sunstein, 2002). The basic principle of toxicology and chemical risk assessment is the dose–response relationship (Benford, 2000; Nau et al., 2002). For PPPs, the following two approaches are used to evaluate safety: hazard-based assessments and risk-based assessments. Hazard-based assessments are used in the approval processes to characterize severe nonthreshold mechanisms, such as carcinogenicity. Risk-based assessments are used in the approval and maximum residue level determination processes to characterize threshold mechanisms. In general, inconsistencies can materialize if both approaches are applied to the same hazard, as is the case in some PPP assessments (Ågerstrand & Beronius, 2016; Barlow et al., 2015; Nordlander et al., 2010). The application of classical risk assessment comprises four steps that generate preliminary forms of knowledge criteria, as elaborated in Table 7 and Table 8.

Table 7: Preliminary forms of knowledge criteria in different steps of risk assessments[15]

Step	Preliminary forms of knowledge criteria
1 Hazard identification	Identification of key negative endpoint (for example, neurotoxic), intrinsic properties of the hazard (acute, short-term, subacute, subchronic, chronic)
2 Hazard characterization	Definition of lethal dose (LOD50) and "no (adverse) effect levels" (NOAEL/NOEL) based on animal data; extrapolation to humans under consideration of safety factors in the form of "acceptable daily intakes" (ADI) for chronic exposure and "acute reference doses" (ARfD) for acute exposure
3 Exposure assessment	Calculation/Estimation of exposure
4 Risk characterization	Aggregation of hazard characterization and exposure assessment

Table 8: Preliminary forms of knowledge criteria in different steps of hazard assessments[15].

Step	Preliminary forms of knowledge criteria
1 Hazard identification	Identification of key negative health endpoints, e.g., carcinogenicity, endocrine disruption potential
2 Hazard characterization	Hazard classification

The outcomes of these processes are the different kinds of knowledge criteria used in determining food safety. Hazard classifications describe the severity of an effect (Henson & Traill, 1993) and the "inherent property of an agent or situation having the potential to cause adverse effects" (Barlow et al., 2015). Hazard classifications are criticized as denying real exposure probabilities (Nordlander et al., 2010). Risk characterizations might describe the levels of risk and probabilities, thresholds, likelihood and severity of an adverse effect or harmful potential and exposition (Barlow et al., 2015; König et al., 2010; Renwick et al., 2003; Whaley & Halsall, 2016). Essential in determining knowledge criteria is the detailed presentation of uncertainties that appear in the various steps determining those criteria (Barlow et al., 2015; Fraiberg & Trebilcock, 1998; Karabelas et al., 2009; Nordlander et al., 2010; Renwick et al., 2003; Tucker, 2008). Human health risk assessments suffer from data that are indirect (from animal studies)

[15] Content of Tables is based on Barlow et al. (2015); Benford (2000); D'Mello (2003); Dreyer and Renn (2009); EC and EP (2005); ECHA (2017); EFSA (2017); EP and EC (2002); Erlacher and Wang (2011); Herges et al. (2017); IoM and NRC (2003); König et al. (2010); Mostafalou and Abdollahi (2013); Nau et al. (2002); Nordlander et al. (2010); Renwick et al. (2003); Stornetta et al. (2015); Szajkowska (2012); Tucker (2008); van der Meulen and van der Velde (2004); Whaley and Halsall (2016).

and must be extrapolated, that are incomplete (not fully understood dose–response relationships), or that present contradicting evidence (Benford, 2000; Fantke et al., 2012; Rhomberg et al., 2013). The concepts of dose–response relationships and thresholds in risk assessment are discussed because they require the exact determination of a threshold, which appears to be difficult given the uncertainties and variabilities (Benford, 2000; Crawford-Brown, 1999; König et al., 2010; Slikker et al., 2004). In addition, in setting threshold levels, multiple exposures in regulations are not taken into account, although cumulative and aggregate mechanisms are highly relevant (Boobis et al., 2008; Kennedy et al., 2015; Reffstrup et al., 2010).

Evaluation

Evaluating knowledge criteria is an important step in the risk-assessment process, which is part of risk management within the regulatory process. Policymakers and regulators function as filters of evidence and decide which forms of societal or scientific information are used in the decision process (Vogel & Delfini, 2008). Because approving PPPs and setting maximum residue levels is highly dependent on the outcomes of scientific risk assessment (EC & EP, 2005b, 2009), regulators have few options in adapting the translated outcomes (regulations) through evaluations based on these classifications or thresholds. For PPPs, an evaluation always relies on data from the risk assessment (Travisi et al., 2006); however, in general, there is little certainty in the translation of health effects into policy (Sexton et al., 2007). Nonetheless, the Sanitary and Phytosanitary Agreement suggests an evaluation beyond the scientific risk assessment in the form of relevant economic factors and trade impacts (WTO, 1995). Furthermore, the multidimensionality of PPPs risks means that important tradeoffs exist between agricultural productivity and safety, which imply a high relevance of socioeconomic evaluations (Florax et al., 2005). These evaluations of PPPs are found in the literature in various but nonharmonized forms (Table 9). Some evaluations are related more to classical risk assessment (e.g., ratings, classifications), mixed forms (e.g., multicriteria decision analysis), economics (cost–benefit analysis), or social sciences (impact assessments).

Table 9: Evaluation methods for PPPs

Evaluation	Specification and Literature
Ratings and Classifications	• Risk ratios, scorings, risk matrices, flow charts, relative risk ranking (Stornetta et al., 2015; Van der Fels-Klerx et al., 2018)
Multi-Criteria Decision Analysis	• Fazil et al., 2008; Van der Fels-Klerx et al., 2018
Risk-Utility evaluations	• Quality of life measures, health-adjusted life years (Cope et al., 2010; Dreyer et al., 2010; Newsome et al., 2009; Stornetta et al., 2015; Van der Fels-Klerx et al., 2018)
Risk-Risk evaluations	• Fraiberg and Trebilcock, 1998; Graham and Wiener, 1995; Hansen et al., 2008; Nordlander et al., 2010
Risk-Benefit evaluations	• Cropper et al., 1992; Starr, 1969
Cost-Benefit evaluations	• Cerroni et al., 2013; Fraiberg and Trebilcock, 1998; Harper and Zilberman, 1992; Sexton et al., 2007 • Marginal Analysis of welfare costs (Lichtenberg et al., 1988) • Benefits through quality improvements (Babcock et al., 1992; Kawasaki & Lichtenberg, 2015; Sexton et al., 2007)
Cost evaluations	• Human Capital Approach, Cost of Illness (Buzby et al., 1998; Caswell, 1998; Henson & Traill, 1993; Van der Fels-Klerx et al., 2018) • Costs for innovation (Nordlander et al., 2010; Sexton et al., 2007) • Cost of resistance (Sexton et al., 2007) • Private costs for farmers (Sexton et al., 2007)
Willingness to Pay/Willingness to Accept	• Caswell, 1998; Cope et al., 2010; Eom, 1994; Florax et al., 2005; Skevas et al., 2013
Impact assessments	• Social impact assessments (Cope et al., 2010; Dreyer et al., 2010; Vanclay, 2002) • Health Impact assessments (Fantke et al., 2012)
Hazard Indices	• Cross and Edwards-Jones, 2011

The literature offers a broad set of tools with which to evaluate safety but suffers from various issues. The main criticisms include complexity of effects that are difficult to illustrate and map in models, limited availability of required data, high effort and costs, high variability of results, and lack of comparable studies because of missing guidelines in PPP risk and hazard impact evaluations (Cope et al., 2010; Fantke et al., 2012; Hansen et al., 2008; Sexton et al., 2007; Skevas et al., 2013; Van der Fels-Klerx et al., 2018).

Value criteria

Value criteria are characterized by a valuing decision on safety based on a previous evaluation. In regulatory practice, we find criteria such as “acceptable” or “tolerable” levels of risk or safety as well as “precautionary” principles. In addition, in socioeconomic literature, we find criteria such as “socially optimum” and “individually optimum” levels.

Acceptability or tolerability describe judgments in risk management pertaining to the acceptability of a risk or hazard in a society. In the current practice of risk analysis within the European context, socioeconomic criteria are not taken into account; only data on the negative impacts on human health based on risk estimates are formally considered (EC & EP, 2002). The issues involved in determining these criteria are that the acceptance of risk must be value-laden and the acceptable level of risk must also have risk management or societal consideration. In general, acceptable levels of risk are difficult to determine due to the heterogeneity in consumer characteristics and preferences (Dreyer & Renn, 2009; Renwick et al., 2003; Shaw, 2005; Wagner, 2016). Regulators are decisive factors in this question because they act as filters of evidence (Vogel & Delfini, 2008). This is problematic because it has been shown that regulators, and thus regulations, are influenced by lobbyism from both the environmental and producer sides (Cropper et al., 1992; Sexton et al., 2007) and that politicians and judges suffer from heuristic biases (Sunstein, 2002).

The precaution criterion is a decision about the status of knowledge. If uncertainties are too high, precautionary bans or measures might be applied to protect public health. These might be applied in cases of incomplete risk assessments and could help avoid unintended tradeoffs in cases of scientific uncertainty (EP & EC, 2002; Hansen et al., 2008; Nordlander et al., 2010). In general, precaution suffers from incoherence, which implies that the application of the criterion is inappropriate in decisions with fatal outcomes (Peterson, 2006, 2017).

Economists introduced the concept of socially optimum levels of safety for PPPs (Henson & Traill, 1993; Sexton et al., 2007). Because there is no realistic scenario in which a society can avoid all hazardous PPPs, safety is determined through optimal levels of safety while acknowledging tradeoffs (Henson & Traill, 1993). The socially optimal level of pesticide use is an outcome of maximizing the net benefit to society, which includes

net benefit to consumers, farmers, producers of chemicals, and the environment (Sexton et al., 2007). The problem with this concept is that it does not consider the distribution of risks in a society (Henson & Traill, 1993).[16]

The individually optimal level is based on a similar argument, with the difference that this optimal safety is determined not for society as a whole but for the individual. It differs from the socially optimum level by not internalizing external factors, such as environmental effects. Issues related to its determination arise through dependence on many factors, *inter alia*, risk perception (Henson & Traill, 1993; Pollak, 1998; Salanié & Treich, 2009; Sexton et al., 2007). In general, market outcomes relate to individual optimum levels because they do not include externalities, such as environmental effects; therefore, markets do not provide a socially optimum level without regulation (Sexton et al., 2007).

Standards

In the last step of risk analysis practice, food safety is defined by fixing legal standards based on value criteria. This step comprises the final decision on food safety, which also implies a decision on which criteria become evident and socially accepted knowledge; therefore, we define standards as the highest level of risk analysis in the framework. We find two types of standards for PPPs: approval or reapproval of PPPs with respect to intentional introduction in the production of agricultural commodities, and maximum residue levels with respect to unintentional presence in the consumption of food (IoM & NRC, 2003; van der Meulen & van der Velde, 2004). Furthermore, we find financial (dis-)incentives in the form of, for example, producer subsidies, which are not further discussed here because they are consequences, not criteria (Sexton et al., 2007).

The approval of PPPs in the EU is regulated by EC 1107/2009. The approval process is divided into the following two procedures: approval of the active substance and approval of the PPP. In terms of active substances, the approval is always limited in time: a maximum of 10 years for the first approval and 15 years for reapproval. Reapproval

[16] A similar concept used to determine safety is provided by Harper and Zilberman (1992): the "Safety minimum standard," which allows a weighting of costs and benefits conditional on some minimal safety standards (Harper & Zilberman, 1992).

enables a new evaluation of the substance based on current scientific and regulatory standards. The approval of PPPs is the responsibility of the European member states but is possible only if all components, including active substances, have been approved. The time limitation depends on the approval of ingredients and active substances (BVL, 2017; Damalas & Eleftherohorinos, 2011; EC & EP, 2009).

Maximum residue levels are regulated by regulation 396/2005 in a harmonized process for all foodstuffs. National risk-assessment authorities, such as the German BfR, estimate maximum residue levels for different kinds of food taking into account PPPs in plants and consumer exposure as it relates to diet. After further EFSA assessment, EU-harmonized maximum residue levels are usually set based on the "As Low As Reasonably Achievable" principle (BVL, 2017). This is a balancing act between production and consumption, with considerations of health, misuse, and trade conditions. The issues in this case might be the heterogeneity of the farms (producers) that apply PPPs, especially in the EU, where harmonized maximum residue levels and highly heterogeneous climatic conditions exist (Maclachlan & Hamilton, 2010; Sexton et al., 2007). National and European food-monitoring programs observe compliance of products with maximum residue levels. If the maximum residue level is exceeded, possible trade restrictions are imposed (BVL, 2017); if it cannot be achieved in practice, PPPs cannot be approved (BVL, 2017).

Standards based on scientific risk analysis always face a fundamental contradiction: science is always understood as an evolving process in which theories are continuously developed and disproved; however, standards are needed to fix a state-of-the-art practice (Wagner, 2016). What is considered to be safe today might be disproved tomorrow by a new or alternative method. Regulations, especially in Europe, acknowledge this issue with time limitations on approvals. This implies that safety is a relative, rather than an absolute, construct. In practice, this is observed with PPPs because their safety depends on the availability of substitutes; if no less-hazardous substances are available, a critical substance might be approved and thus considered as safe (Storck et al., 2017). In addition, standards must fulfill two expectations: first and foremost, protect consumers in the best possible way, and second, avoid a possible function as a trade barrier. Thus, the harmonization of standards is necessary to maintaining trade. In many cases, it is difficult to decide whether a deviation from internationally accepted standards is justified based

on consumer protection; therefore, scientific risk assessment is set as the basis for these decisions. Various authors from jurisprudence deal with the question of the extent to which science is required by the Sanitary and Phytosanitary Agreement in international trade because it remains unclear to what degree risk assessment or precautionary reasons are decisive (Cunningham, 2005; Gruszczynski, 2007; Kerr, 2009; Wagner, 2016; Walker, 2001; WTO, 1995).

The EU's PPP regulations do not take into account socioeconomic criteria in deciding on legally fixed safety. Regulations formally rely purely on data about the negative impacts on human health based on risk estimates (Cope et al., 2010; König et al., 2010; Reinert, 2015; Verstraete, 2014). This might be problematic, as different levels of PPP regulations have a significant impact on production and trade[17] and must address wider societal concerns. In general, PPP regulations face the issue that food safety problems can easily spread globally as a result of trade and require a uniform policy; however, the use of PPPs and consumption is local and heterogeneous (Sexton et al., 2007). In the EU, PPPs are regulated within a complex regulatory landscape comprising various national and supranational institutions and processes that increase the likelihood of inconsistencies (EC & EP, 2005b, 2009; ECHA, 2017; Handford et al., 2015; Rhomberg et al., 2013; Storck et al., 2017).

PPP bans can be problematic because they decrease the number of available active substances and increase a buildup of resistance (Karabelas et al., 2009; Sexton et al., 2007) or the application of even worse substitutes (Nordlander et al., 2010). Policy often ignores the possibility of risk-reduction measures (e.g., application and drainage requirements to reduce risk) and bans pesticides without considering their economic benefits (Sexton et al., 2007).

3.1.4.3. Evidence practice of consumers

Risk analysis in food safety is a highly institutionalized evidence practice. While standardized risk assessments attempt to achieve an objective and fair safety determination, risk management must negotiate various interests, such as consumer safety

[17] The literature reports mixed results regarding the effects of PPP regulations on trade (Disdier et al., 2008; Drogué & Demaria, 2012; Essaji, 2008; Handford et al., 2015; J. S. Wilson & Otsuki, 2004).

and trade. The second relevant process in determining food safety is the evidence practice by consumers, which is more vague. In line with Galison (2010), we considered this not as a reflection or interpretation of the risk-assessment process but as one of self-reliance. This differs from recent definitions describing consumer evidence practice as, for example, risk perception and "a critical interface between scientific facts and personal opinions and values" (Ropeik, 2011). In the literature, we find various terms for consumer practice, such as intuitive toxicology (Kraus et al., 1992), risk perception (Slovic, 1987), people's risk assessment (Bieberstein & Roosen, 2015), risk evaluation (Bouyer et al., 2001), subjective food safety (Grunert, 2005), and qualitative criteria (Henson & Traill, 1993).

This indicates two things. First and foremost, the literature recognizes that consumers use their own determination processes to evaluate food safety and risks. Second, these differences in definitions show inconsistencies in the understanding of the consumer evaluation process. In the literature on food quality, food safety is described as a credence attribute because consumers are not able to directly evaluate the safety of the offered products before and after purchasing (Grunert, 2005). This is especially true for PPPs because their effects are mostly chronic and delayed; therefore, they cannot be directly related to a specific consumption event (Henson & Traill, 1993; Shaw, 2005). Nevertheless, consumers have their own understanding and definition of food safety independent of single purchases, which differs from an understanding of the definition in risk analysis. In the following section, we collect different expressions of safety as found in research and classify them in the criteria structure. Contrary to risk analysis, it is not possible to separate criteria on an institutional basis because all safety-determination criteria are used by the consumer. Additionally, it is difficult to separate consumer criteria conceptually because the entire safety determination process is inclusive and interconnected. Nevertheless, we decided to split them because it is common in the recent literature to differentiate between "objective" and "subjective" elements in safety determination. The terms used in our framework follow the classifications in the risk analysis process: (1) knowledge criteria, (2) value criteria, and (3) behavior, as summarized in Figure 7. Between levels (1) and (2), we place the perception process.

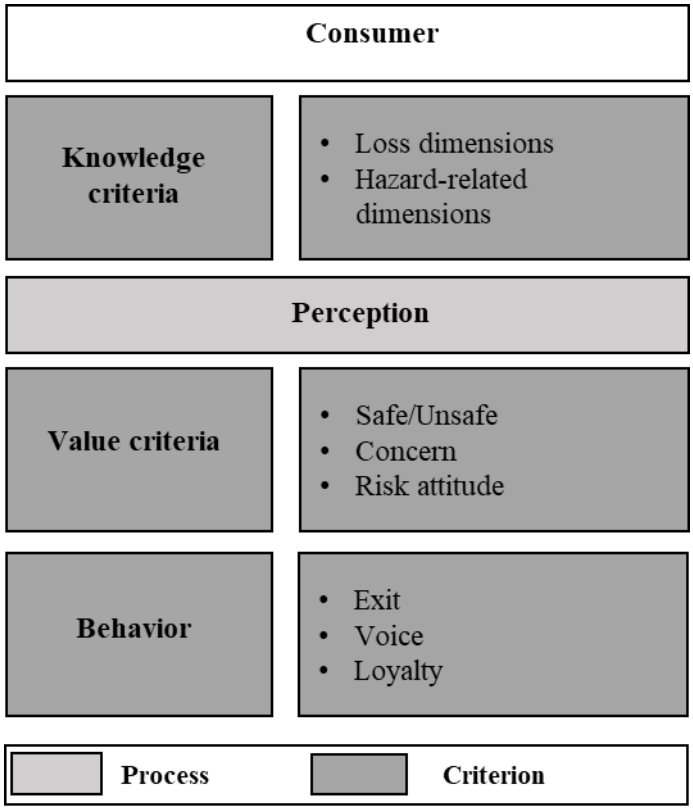


Figure 7: Food safety criteria in consumer evaluation

Knowledge criteria

Following the classification of risk analysis criteria, knowledge criteria represent the relevant underlying dimensions used by consumers for safety determination. As this classification has not been used systematically in the literature thus far, we determined the following conditions for their characterization: All concepts describing relevant safety dimensions for consumers and their reasonable determination require some kind of knowledge and could be determined objectively using science or scientific methods. In addition, they should fulfill the general definition of criteria, which are required to be the basis for decisions or judgments. Because we do not claim that consumers are able to determine these criteria in the form of unbiased quantifications or numerical expressions, these criteria are seen as non-numeric, qualitative descriptions. The various concepts describing factual dimensions in consumer safety determination are elaborated below.

Personal loss dimensions

The first type of knowledge criteria describes the types of losses that are relevant to consumers and stem from a concept introduced by Yeung and Morris (2001), who

determined six individual loss dimensions to predict overall risk, specifically food risks based on the components of perceived risk in product purchase in Kaplan, Syzbillo, and Jacoby (1974). The first dimension, "physical losses", understood as a safety criterion, describes negative health effects. In the case of PPPs, this might include acute intoxications or chronic diseases (Yeung & Morris, 2001). In other research contexts, this criterion is considered a "fatality of consequences", including acute or chronic effects and type of death (Fife-Schaw & Rowe, 1996; Mullet et al., 1993; Slovic, 1987; Sparks & Shepherd, 1994). Second, "psychological losses" describe food-risk concerns, for example concerns about a PPP when buying a certain product. Third, "financial losses" describe possible replacement costs for risky foods, the cost of medications and income losses while recovering from the adverse effects of a food containing hazardous substances. Fourth, "time losses" refer to the time wasted in replacing affected food or in illness as a result of consuming a PPP-contaminated product. Fifth, "performance losses" describe the possible adverse effects of the hazard on taste or nutritional value, and sixth, "social losses," are the social embarrassments resulting from the use of contaminated food products (Kaplan et al., 1974; Yeung & Morris, 2001). The authors describe the theoretical concept as suitable for all types of food risks. To the best of our knowledge, there is no application of these loss categories in PPP safety determination[18]; nevertheless, the concept offers the possibility of describing relevant effects for consumers, including various health effects and wider financial, social, and psychological concerns.

Hazard-related dimensions

The second type of knowledge criteria stems from the psychometric paradigm invented by Slovic (1987). The concept describes two hazard-related dimensions in consumers' risk determination: the "unknown" and the "dread" factors, both of which comprise multiple items that include relevant safety dimensions for consumers.

The unknown factor comprises a "lack of observability", whether the hazard is "known to the people exposed" and "known to scientists", whether it is a "new or known risk"

[18] Studies on other food hazards have shown good performance in predicting overall risk perception and significant differences among the loss categories, which might support the argument to include these loss dimensions in food safety determination (Hornibrook et al., 2005; Mahon & Cowan, 2004; Yeung & Yee, 2002).

and whether "effects are delayed or immediate". In different studies, PPPs score low to middle on the unknown factor. The second factor in the psychometric paradigm is the dread factor, which includes the "potential for a global catastrophe," the "level of control of the hazard" (including the possibility of reducing it) and the "general development of risk" (either decreased or increased) of a PPP. PPPs score medium to high on the dread factor (Fife-Schaw & Rowe, 1996; Slovic, 1987; Sparks & Shepherd, 1994).

Classical psychometric paradigm studies indeed show high explanatory power but use mainly aggregated data, which ignores individual differences among consumers. This has been criticized because individual differences have been shown to be highly important in explaining risk perception. The inclusion of individual data lowers, but does not neglect, the relevance of hazard-related dimensions (Bronfman et al., 2008; Marris et al., 1997; Siegrist, Keller, et al., 2005).[19] Additionally, the concept ignores the socio-political conditions which influence public opinion on hazards (Boholm, 1998). Nevertheless, the psychometric concept describes a validated set of relevant dimensions determining safety that are characterized as knowledge criteria in this study.

Knowledge criteria as food safety criteria

In both concepts, the individual and the hazard-related, we find various, differentiated dimensions that play a role in determining food safety in relation to PPPs and can function as food safety criteria. They include direct health effects but also concerns for indirect and wider effects on individuals or society. Thus, they confirm that public safety determination is not based on a reflection of statistical numbers. The consumer knowledge criteria concepts confirm the influence of qualitative understandings of hazards such as PPPs (Boholm, 1998). However, few concepts are specified at the level of the relevant underlying safety dimensions for specific hazards. In existing studies, it is unclear how dimensions of food safety are derived, indicating that there might be dimensions missing because they are outside the concept frames. The previously described concepts partly overlap and we can categorize most of the dimensions of

[19] Research on individual differences is elaborated in the next section on perception.

consumer knowledge criteria content-wise using the basic qualitative structure of risk analysis knowledge criteria, hazard identification and exposure.

In risk analysis, hazard identification is defined as follows: “the identification of the type and nature of adverse effects that an agent has an inherent capacity to cause in an organism, system, or (sub)population” (Barlow et al., 2015). Among consumer knowledge criteria concepts, we find criteria such as health effects and indirect effects such as financial or social losses, primarily in the concept of loss dimensions. This is different to knowledge criteria from risk analysis, which consider exclusively the direct negative health effects of PPPs (König et al., 2010).

In risk analysis, exposure is defined as the “evaluation of the exposure of an organism, system, or (sub)population to an agent (and its derivatives)” (Barlow et al., 2015). In consumer safety determination, the following related criteria can be found mainly in the psychometric paradigm but also in other studies: concern for future generations (Miles & Frewer, 2001; Slovic, 1987), concern for vulnerable groups (Miles & Frewer, 2001), and the distribution and equity of risks and benefits (Dreyer et al., 2010; Fischhoff et al., 1978; Slovic, 1987).

Additionally, we find other relevant dimensions that go beyond the risk analysis knowledge criteria frame, for example, the possibility of reducing a risk or the risk trend (decreasing or increasing).

The authors of the original concepts, Yeung, Morris and Slovic, see the dimensions as being mostly relevant in effective policy or communication. Comparison with risk analysis knowledge criteria shows that the qualitative dimensions can also be useful during the risk-assessment stage or in framing risk analysis. Underlying dimensions used as safety criteria could also be relevant in the conceptual definition of socially acceptable safety levels (Aven, 2018; Klinke & Renn, 2002).

Perception

In the context of our framework, risk perception is the evaluation process located between the knowledge criteria and the value criteria. This follows the argument that perception might lead to the social amplification of risk and that risk perception influences risk attitudes and thus indirectly influences behavior (Kasperson et al., 1988; Lobb et al., 2007). As described in the preceding sections, it is impossible to draw a clear conceptual

line between knowledge criteria and perception factors; therefore, perception includes all factors not clearly classifiable as knowledge criteria. Most studies do not measure how consumers determine safety with respect to risk but rather how different, non–rational factors influence the safety–risk evaluation; thus, perception is not a separate criterion because it does not describe a determination of safety but rather a process that influences the determination of safety and that is influenced by a variety of factors. These factors are identified in the following section and might influence or bias an individual evaluation of safety and explain interindividual or group differences in the risk perception of PPPs. These factors might explain why different consumer groups are concerned about different types of food risks (Cunha et al., 2010; Roosen et al., 2005), or why there exist controversial opinions on food safety in terms of PPPs.

Following Bouyer et al. (2001), we classified the influencing factors in the "perceiver-linked" and "hazard-linked" factors. In addition, we found "external factors." Because the literature on influencing factors is very extensive but not the focus of this paper, these factors are not discussed in detail. In addition, because this paper does not aim to provide a quantitative meta-analysis, comparability of study designs and the strengths of effects (positive or negative) are not discussed.

Perceiver-linked factors

In addition to physiological processes, such as brain mechanisms, the perception process can be influenced by various factors, which are summarized in Table 10. We categorized "perceiver-linked" factors into psychological factors, sociodemographic determinants, and sociocultural determinants.

Table 10: Psychological factors in the explanation of risk perception.

Factors	Specification and Literature
Psychological factors	
Bounded rationality*	• Examples: Framing or subconscious mental shortcuts, specifically optimism bias or loss aversion (Ropeik, 2011; Sunstein, 2002)
Affective heuristics*	• Representativeness, availability and anchoring comprise simplifying mechanisms to evaluate risks (Kasperson et al., 1988) • Strong drivers of intuitive judgments • Might lead to systematic errors of risk estimates (Sunstein, 2002)
Personality factors*	• Level of anxiety (Bouyer et al., 2001)
Emotions*	• Used by various authors in different theoretical concepts (for example (Loewenstein et al., 2001; Peters et al., 2004)) • Emotions connected to outcome seem to be more important than probability (Sunstein, 2002)
Sociodemographic factors	
Gender	• Bieberstein and Roosen, 2015; Bouyer et al., 2001; Byrne et al., 1991; Dosman et al., 2001; Dunlap and Beus, 1992; Knight and Warland, 2004; Miles et al., 2004; Nayga, 1996; Siegrist et al., 2005; Verbeke et al., 2005
Age	• Bouyer et al., 2001; Dosman et al., 2001; Dunlap and Beus, 1992; Knight and Warland, 2004; Lin, 1995; Miles et al., 2004; Nayga, 1996; Siegrist et al., 2005; Verbeke et al., 2005
Education	• Byrne et al., 1991; Nayga, 1996; Verbeke et al., 2005
Household income	• Byrne et al., 1991; Dosman et al., 2001; Miles et al., 2004
Number of children	• Dosman et al., 2001
Sociocultural factors	
Values	• Generally important in explanation of risk perception (e.g., Dreyer et al., 2010; Hansen et al., 2003) • Self-centered and altruistic values (Bieberstein & Roosen, 2015) • Group or cultural values, cultural attitudes, worldviews, and political orientation (Bouyer et al., 2001; Douglas & Wildavsky, 1983; J. Hansen et al., 2003; Kahan, 2016; Kahan et al., 2007; Kasperson et al., 1988; Peters et al., 2004; Ropeik, 2011)
Normative heuristics*	• Shaped by the normative concept or accepted societal rules[20] (Sunstein, 2002)
Cultural handling of food*	• Right or wrong, based on traditional processes whose importance appears to be decreasing (Lusk, 2013; Pollan, 2006; Spiekermann, 2011)

* Factor is a general concept in risk perception but not proven explicitly for PPPs, to the best of our knowledge.

[20] An example would be the neglect of risk-benefit estimations because calculating deaths is not accepted in cases of human harm (Sunstein, 2002).

Table 10 shows that a great variety of individual factors has an influence on the food safety determination process, which might lead to under- or overestimating the risk of a PPP and thus influence safety determination. First, psychological phenomena might cause systematic errors in the determination of safety. Second, individual characteristics influence how people determine safety; how hazardous PPPs in food are perceived depends to some extent on age, gender or education. Third, some research aims to explain risk perception differences using sociocultural factors. For PPP safety determination, this means that the value orientation and normative context of a person or a group influence the outcome of the safety evaluation. Safety concerns regarding PPPs and the acceptance of safety-related information might be motivated through political intention, or as a part of identity-building processes (Kahan, 2016). It is important to keep these various factors in mind when assessing public safety determination. Doing to may help one avoid misinterpretations or explain why certain consumer groups are dissatisfied with safety standards while others are not at all concerned. This is especially the case when interpreting the value criteria, elaborated in a later section.

Hazard-linked factors

In addition to perceiver-linked factors, risk perception is influenced by nonrational hazard-linked factors connected with a hazard's intrinsic characteristics. Research finds a numerical overestimation of risk from PPPs (Williams & Hammitt, 2001). Different approaches explain why some hazards are generally overestimated when compared to statistically derived numbers in public safety perception or to other hazards. These factors include the "nature of the hazard." The type of hazard affects perception (Dosman et al., 2001; Kher et al., 2013). The literature specifically identifies differences between natural and technical hazards (Fischhoff et al., 1978; Kaptan et al., 2017; Kraus et al., 1992; Ropeik, 2011; Sjoberg, 2000; Williams & Hammitt, 2001), and technological or lifestyle issues (Miles et al., 2004). PPPs are a technological hazard in foods and are therefore likely to be overestimated compared to other food hazards. Related to this idea is the phenomenon of the "stigmatization" of, for an example, a technology. Stigmatization might be an additional mechanism whereby perceived risk is amplified in certain cases (Kasperson et al., 1988); specific hazards or PPP substances might be stigmatized. Another important influencing factor is weighing the "perceived risks and benefits" associated with the substance of interest (Alhakami & Slovic, 1994; Finucane et al., 2000;

Miles & Frewer, 2001; Peters et al., 2004; Ropeik, 2011; Siegrist et al., 2000; Starr, 1969; Sunstein, 2002; Verbeke et al., 2005, 2007). Here, the relationship is reversed: if high risks are associated, perceived benefits are reduced and vice versa (Alhakami & Slovic, 1994; Finucane et al., 2000; Peters et al., 2004). It has been shown that certain consumer types acknowledge the benefits of PPPs (Saba & Messina, 2003), but generally the risks of pesticides are rated higher than their benefits (Alhakami & Slovic, 1994). It appears that safety concerns related to PPPs are rather high unless PPPs bring about significant benefits in agricultural production, food safety, or food security (Sexton et al., 2007). The inverse risk-benefit perception might contribute to overestimating the risks and, vice versa, underestimating the benefits of PPPs. Additionally, the mechanism might increase perceived risk if the focus of public communication and discourse is predominantly on risks rather than benefits (Kaptan et al., 2017). All these hazard-linked factors indicate that PPP risks are, in terms of their characteristics, likely to be overestimated in food safety determination. This should be kept in mind if consumer food safety criteria are used in decision-making processes.

External Factors

External factors influence the risk perception process but are not related to the hazard or the perceiver. They are somewhat situational and depend on external influences. Under external factors, we include "food scandals", which might amplify the risk perception of related hazards during certain time periods (Kasperson et al., 1988; Lobb et al., 2007; Rieger et al., 2017; Verbeke et al., 2007). Additionally, safety perception is influenced by "the general situation of traceability and recall" (Hobbs et al., 2005; Kher et al., 2013; Van Rijswijk & Frewer, 2008), "the amount of available information" (Buzby et al., 1998; Eom, 1994; Finucane et al., 2000; Henson & Traill, 1993; Peters et al., 2004; Travisi et al., 2006) and "media influence" (Mccluskey & Swinnen, 2004; Sunstein, 2002; Verbeke et al., 2007). "Trust" in various situations was shown to be an important influencing factor on risk perception but its influence depends on the related institution or stakeholder (Chryssochoidis et al., 2009; Kaptan et al., 2017; Lobb et al., 2007; Ropeik, 2011; Siegrist et al., 2000; Siegrist, Gutscher, et al., 2005; Williams & Hammitt, 2001).[21] The preceding

[21] For example, trust in EU institutions as an information source varies between 88% and 40% across EU member states (EFSA, 2019).

list of factors indicates that the evaluation of public safety determination in the case of PPPs must take into account related framework conditions independent of the hazard itself.

Value criteria

Equivalent to the risk analysis process, there is a level of "value criteria" for consumers to determine safety. "Value criteria" comprise a valuing decision about safety that might be determined based on or influenced by "knowledge criteria" or "perception". The literature mentions three types of value criteria: first, the stated claim that something is or is not safe (Leiss & Powell, 2004), and second, stated concerns about PPPs (Bruhn et al., 1992; Byrne et al., 1991; Dunlap & Beus, 1992; Knight & Warland, 2004; Miles et al., 2004; Nayga, 1996). Often levels of concern are measured in numerical, unidimensional scales. Concerns with PPP risks appear to be highly rated compared to other food risks (BfR, 2016; EFSA, 2010).[22] Third, various scholars describe risk attitudes in a way that allows one to classify them as value criteria. For example, Pennings et al. (2002) describe attitudes as "the decision-maker's interpretation of content of the risk and how much (s)he dislikes the risk"; Wilcock et al. (2004) describe attitudes as "permanent and stable evaluative summaries". Attitudes have been shown to influence intentional behavior. This relationship links value criteria and behavior, which are considered the final step in determining safety in consumer evidence practice.

Value criteria as food safety criteria

Value criteria in the form of closed survey questions or item batteries are frequently used in studies as dependent variables in risk perception studies or to measure public safety determination. The literature comprises a broad base of validated instruments and empirical evidence. These carry some weight in discussions of public safety determination and are used in official EU and national consumer surveys (BfR, 2016; EFSA, 2010); nevertheless they must also be examined critically. First, they must always be seen as depending on the many influencing factors described in the preceding sections. Second, they are dependent on survey instruments. For example, studies show

[22] These results are also confirmed by a more recent version of the Eurobarometer published after the literature search (EFSA, 2019).

overestimations of food risks in closed survey questions. The proportion of respondents expressing worries about chemical contaminants in closed questions is 85%. In open questions, 17% of respondents mention concerns (Gaskell et al., 2017). Third, they can provide no information on the underlying reasons for concerns or decisions about safety or about actual behavior, but rather help to identify critical issues.

Behavior

The last level in the consumer determination of safety is the actual behavior of the consumer. This must be differentiated from "value criteria" because, for example, "different attitudes do not necessarily lead to behaviors that increase the safety of the food consumed" (Wilcock et al., 2004). Numerous studies have attempted to explain, *inter alia*, the relationship between consumer attitudes and behavior. These relationships are not elaborated in this study because our aim was to examine different types of food safety criteria. In the case of PPPs, handling or cooking practices are not effective behavior options to improve safety; therefore, there are limited behavioral options to determine safety. Generally, consumers can use various strategies to respond to inadequate policy and/or safety supply only in cases in which quality failures exceed tolerance levels (Hirschman, 1971). According to Hirschman, these possible strategies are "exit", "voice" or "loyalty". Exit strategies might include stopping, reducing, or shifting consumption from one product to another (Roselius, 1971). For PPPs, the exit option is difficult to put into practice because most available products on the market are conventionally produced. An option might be to stop or reduce the purchase of conventional food and to switch to organic products produced without synthetic PPP applications.[23] Existing studies indicate a reduction in purchases due to concerns about PPP residues in fruits and vegetables (Bruhn & Schutz, 1999; Unusan, 2007) and in meat and meat products (Unusan, 2007).[24] The "voice" option includes expressing dissatisfaction directly to the producer or as a general protest (Hirschman, 1971). Here, consumers face different barriers in the case of PPPs. As safety concerns are not related to a specific food product but instead to a PPP,

[23] Organic food products may reduce the intake of PPP residues but have been shown to be partially contaminated (Smith-Spangler et al., 2012).

[24] Additionally, the Eurobarometer, published after the literature search, indicates that European consumers react differently to information on food risks in general: some indicate changes in their behavior while others do not, although the latter remain concerned (EFSA, 2019).

the recipient of complains is not easily identifiable. Information on risk assessment and underlying studies are partially confidential and not available to the public (EC & EP, 2009; Hirschman, 1971; Rosman, 1993). Additionally, consumers require the technical knowledge and financial resources necessary to effectively understand and interact with the complex process of risk analysis; for example, in public consultation rounds in the approval process. Such resources are available to interest groups rather than to individuals (Sunstein, 2018). The voice option is more costly than exiting, especially in cases involving a large number of affected products (Hirschman, 1971). This might be a reason why (successful) complaints of consumers and citizens appear in the form of (publicly supported) third-party actions. Successful examples of the voice option in the US include the ban on DDT or Alar initiated by public complaints (Rosman, 1993). A more recent example from the EU is the initiative "ban glyphosate and protect the people and environment from toxic pesticides" (EC, 2017). The efforts surrounding initiatives are high and therefore only appear in critical cases (Rosman, 1993). The "loyalty" option would be to continue to purchase "and absorb the unresolved risk, indicating that the perceived risk associated with a particular product is tolerable and no greater than alternatives" (Roselius, 1971). This implies that the final consumer determination of safety is highly dependent on purchasing alternatives but also on tolerance levels.

Behavior as a food safety criterion

In general, behavior is relatively easy to assess. It is possible to analyze decreasing sales or the number and types of complains. Exit and Voice are in general good instruments with which to express dissatisfaction, but managers need time to react and to adapt the system accordingly. Behavior can have a rather destructive potential in critical cases (Hirschman, 1971); a prominent example is the Bovine Spongiform Encephalopathy crisis in Europe in the 1990s, which entailed high costs for the food system (Leiss & Powell, 2004). For the case-study of PPPs, the preceding arguments show that it is difficult for consumers to express their concern with safety violations through changes in behavior. The barriers and costs involved in complaining or choosing other options are high. Therefore, purchases might not be a suitable food safety criterion in dealing with PPPs but might still be relevant to other cases of safety determination. With PPPs, it is important to be aware of the behavioral options consumers have and the barriers they

face. Related organizations must find solutions and reduce costs to enable consumers to express dissatisfaction (Hirschman 1971).

3.1.5. Discussion

The aim of this study was to provide a conceptual understanding of the complex determination processes in food safety. The developed framework shows that a generalized definition of safe food, as evidenced by socially accepted knowledge, remains impossible. Food safety cannot be described merely as the absence of a harmful substance (Herges et al., 2017) or the inverse of risk (Henson & Traill, 1993). Conceptually, it is determined instead by various criteria at different levels in the risk analysis process or by consumers and also requires valuing decisions. This contradicts recent studies that assume the existence of one well-defined level of safety that can be determined empirically and objectively but might be perceived or interpreted differently. Our framework introduces the issue of determining safety and illustrates how regulators can include science-based thresholds, a socioeconomic optimum, or a safety determination made by the consumer.

The case of PPPs offers interesting insights into safety determination processes because absolute safety in the form of the absence of harmful substances is unrealistic in today's agriculture and nutrition (Babcock et al., 1992; Henson & Traill, 1993; Kawasaki & Lichtenberg, 2015; Sexton et al., 2007). Moreover, Galison's trading zone concept related to evidence practices offers a fruitful basis from which to explain and categorize the process of determining food safety. It provides a different perspective on consumers' determination practice and might also be an interesting approach for structuring other complex social negotiation processes. A characteristic of the trading zone concept is that it can uncover misunderstandings in the exchange of criteria because the concepts themselves, but not necessarily their meanings, are shared among the different zones and actors.

There were limitations to this study that must be considered when interpreting its results. Because the framework was developed based on the literature, which was unable to resolve biased interpretations, the study might suffer from a confirmation bias. To reduce this risk, a systematic literature review and discussions with uninvolved researchers were conducted. The validation procedure could also be a limitation. Following the validation

procedure suggested by Jabareen (2009), we discussed the results with experienced researchers within different conference settings. This was problematic because it was not possible to select the audience and researchers attending the presentations; therefore, it might be that, as opposed to an expert workshop, the participating audience had no experience of the topic being discussed. Because this is a somewhat weak and nonspecific validation method, various arguments support the framework's validity. One argument for its credibility or internal validity was the use of studies from different perspectives and disciplines. Transferability or external validity is not the aim of qualitative research and the justification of transfer is not the responsibility of the developing, but rather of the applying, researcher (Bitsch, 2005). Another limitation might be the topic's high complexity; thus, we assume that the literature examined is incomplete. Because this is not a quantitative meta-analysis, we justify our approach with reference to the concept of theoretical saturation. It is possible for researchers applying this conceptual framework to extend or adapt it with their own classification schemes.

The framework points out the relevance of the differences and peculiarities among the criteria involved in evidence practices. While criteria in risk analysis are determined via a highly institutionalized process and are mostly consistent with each other, consumers' safety evaluations are more complex, interconnected, and hardly tangible. Conceptual and semantic differences among safety determinations lead to various issues and related implications in the societal discourse on food safety.

First, we observe safety determination by consumers as being inconsistent in itself: consumers do not show noticeable forms of "exit" or "voice" at a behavioral level in the general evaluations of PPPs; however, this does not appear to be the case at the value level. Here, we observed a high level of concern about PPPs (BfR, 2016; EFSA, 2010), even if the safety level is deemed acceptable by conformity with regulations. This raises the question as to when inadequate safety exceeds tolerance levels and leads to a change in behavior and what barriers for behavioral options exist (Hirschman, 1971). This question is of major economic relevance because changes in behavior caused sudden economic losses and decreases in consumer trust in previous food scandals, as seen in the aftermath of the Alar scare related to apples in the US in the 1980s (Syddell, 1990) or the Bovine Spongiform Encephalopathy crisis in the 1990s in the EU (Leiss & Powell, 2004;

Raude et al., 2005; Vos, 2000). Therefore, an interesting area of research involves determining tolerance thresholds in food safety.

Second, disregarding or overprotecting consumers can lead to under- or overregulating PPPs (Fraiberg & Trebilcock, 1998). This raises the question of how regulators integrate evidence practices by consumers in their processes, a question discussed in the existing literature (e.g., Salanié & Treich, 2009). We aim to enrich this discussion by identifying different types of criteria. This study shows that numerous factors might influence and bias individual safety evaluations; therefore, the currently used measurements of stated perceptions or concerns might be unsuitable for determining rational and evidence-based safety. An alternative approach is recognizing relevant underlying knowledge criteria, which might be assessed in scientific risk assessments and extend the current dimension of direct negative health effects. One relevant example could be equality of risk (EC & EP, 2005b, 2009; König et al., 2010). In this case, the weak scientific basis for these criteria related to food safety and PPPs is problematic, which might be an interesting area for future research. Regarding behavior, the last level of safety determination, one can argue provocatively whether we need to consider stated concerns as long as consumers do not actively react, as this is their final decision about safety (Finn & Louviere, 1992). Hence it is important to be aware of existing barriers in order to interpret consumer behavior. We emphasize that regulators and policy consultants must recognize the existence of consumers' different levels of safety determination and decide on which level to base their measurements.

Third, there is the question of how to determine the value criteria in the risk analysis process and how to define societal acceptability. As with many risks in modern societies, PPPs always result in tradeoffs in their applications, which are relevant at different policy levels (Beck, 1986; Graham & Wiener, 1995; Sexton et al., 2007). In addition, PPP regulations are directly linked to the continued functioning of free trade; therefore, non-optimal regulations of PPPs as food hazards might threaten consumer safety, societal welfare, or food security (Handford et al., 2015; Wagner, 2016; WTO, 1995). This implies the need for the integration of socioeconomic evaluations, at least in the process of determining value criteria; therefore, transparent and harmonized evaluation methods for socioeconomic impacts, which thus far do not exist, are necessary. In addition, these evaluation methods neglect the distribution of risks and benefits, but these should be

integrated because they appear to be relevant to consumer evaluations. It should be acknowledged that in evaluations optimal levels of safety will most likely vary among different stakeholders depending to what degree the complex externalities of PPPs are internalized (Henson & Traill, 1993; Sexton et al., 2007). This might also be important for risk communication and to justify regulatory measures. If a communicator or regulator knows the actor involved and his/her specific individual optimum of safety, he/she might be able to explain deviations from this optimum with the internalization of further socially relevant externalities. Similarly, risk regulators can also consider the internalization of externalities deemed relevant by other stakeholders to achieve a socially acceptable optimal level of safety.

Fourth, in the process of risk analysis, it should be recognized that the selection of knowledge criteria is contested and cannot be deemed entirely objective. Knowledge criteria provide a solid and comparable basis for safety evaluations but also exhibit weaknesses. As shown in section 3.1.4.2, the main issues are the critical evaluation of available knowledge (Douglas & Wildavsky, 1983; Dreyer & Renn, 2009) and the Weight of Evidence harmonization (Ågerstrand & Beronius, 2016; Rhomberg et al., 2013; Weed, 2005). Furthermore, determining food safety with regard to PPPs is never conclusive and depends on the state-of-the-art practice in science, which is addressed by time-limited approvals (Erlacher & Wang, 2011). This illustrates how difficult it is to formulate an ultimate definition of food safety.

Fifth, a vacuum between the determinations of safety by risk analysis and by consumers can develop that would provide room for communication among nongovernmental interest groups, the media, or private companies.[25] In a worst-case scenario, this might generate social panic, which often does not even need to be connected to a change in the underlying risk (Leiss & Powell, 2004; Loewenstein et al., 2001; Verbeke et al., 2007); consequently, purely science-based risk assessments may be in danger of becoming the subject of instrumentalization. For example, interest groups can neglect or emphasize aspects of scientific uncertainty to support their own agenda (Jasanoff, 1990). In these

[25] An example is the German food retailer ALDI SÜD, which determines its own PPP residue standards for fruits and vegetables at a stricter level than the regulatory requirements (Mempel, 2015).

cases, well-founded decisions based on socioeconomic analysis and consumer determination might be easier to justify.

Based on the above arguments, we emphasize that in food safety discussions, participating individuals should recognize the existence of different levels of safety and their related issues in order to improve public discourse and optimize safety determination. This might lead to more consistent regulations, an increase in trust, and mutual understanding.

3.1.6. Conclusion

Overall, this study shows that food safety, as an issue of scientific, economic, and societal relevance, cannot be determined at one generally accepted level. Rather, a common understanding of the issue is the result of complex negotiation processes. The framework developed here provides an overview of the influencing factors and issues in this negotiation process, and highlights semantic and conceptual inconsistencies in the literature. This helps one understand this complex social issue and such related phenomena as public discourses. The framework integrates various types of knowledge in a conceptual manner, which is a major task of risk research (Aven, 2018; Jabareen, 2009). It offers the possibility of reassessing, contextualizing or generating hypotheses for empirical research in the individual disciplines. In addition, it raises our awareness of how food safety is determined at different levels, which leads to the question of the level at which safety should be based. In risk research and analysis, which criteria should be used to determine acceptable risk or safety is a decisive and nonobvious question. Conceptual research can help answer this question by providing an overview of how existing criteria and issues are determined and by placing them in context (Aven, 2018). In this study, the framework illustrates the difficulties involved in existing risk analysis procedures and the possible sources of public dissatisfaction. This may help foster an understanding among different social groups and also in research, or to justify a reconsideration of currently used criteria in food safety. The framework might provide interesting approaches to solutions to existing issues and ideas for future research, either in the risk-assessment process, the field of safety evaluation by consumers, or in their interactions. The conclusions reached might be very different for each discipline – a critical consideration of scientifically based criteria or a more differentiated view of the

criteria used by consumers. The study further shows the potential of applying conceptual framework research to risk research, which is often characterized by multidisciplinary and complex areas of interest.

3.1.7. Acknowledgements

This article is based on research supported by funding from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Research Group 2448. We thank Professor Ellen Goddard for her helpful and enlightening comments in the stage of the framework development. We greatfully acknowledge the anonymous reviewers and the editor for their constructive comments on earlier drafts of the manuscript.

3.2. Which criteria do consumers use to evaluate the safety of food?

Publication: Hassauer, C., & Roosen, J. (2019). Which criteria do consumers use to evaluate safe food? *Selected Paper prepared for presentation at the 2019 Agricultural & Applied Economics Association Annual Meeting, Atlanta, GA, July 21 – July 23*

3.2.1. Abstract

For regulators it is challenging to consider food safety criteria that are important for consumers without over- or underestimating them. An inductive approach was used to provide a qualitative description of used food safety criteria. Fifty-three in-depth interviews were analyzed using a grounded-theory approach. Results show the possibility to differentiate between criteria that describe the constitution and the evaluation of safety. This provides an approach to analyze consumer framing of food safety, and possibilities to optimize case-specific risk communication in critical food-safety issues.

3.2.2. Introduction

The determination of optimal and acceptable levels of food safety requires a complex negotiation of various interests, including those of consumer protection, economic considerations, and societal factors. Therefore, it is a pertinent question how consumers frame food-safety determinations and which criteria enter, and are relevant in, the food safety discourse. The consideration of consumer safety determination becomes particularly important for uncertain or ambiguous risk phenomena. These can usually not be regulated solely based on risk assessments because they may entail normative, evaluative or cognitive conflicts (Klinke & Renn, 2002). The relevance of acknowledging the public in food safety determination in addition to classical, scientific risk assessment is acknowledged in politics as well as in the scientific literature. For example, the European Food Safety Authority stated the prioritization “of public engagement in the process of scientific assessment” as their first strategic objective in the 2020 Strategy (EFSA, 2015a). Scholars established suggestions for new approaches in risk analysis in food safety to overcome the issue of the integration of societal factors as public safety determination in the current system, for example in risk prioritization or evaluation

(Dreyer & Renn, 2009; König et al., 2010; Ruzante et al., 2010). This is consistent with the understanding of the concepts of acceptable risk and safety as societal constructs. These require the inclusion of actors for determination, goes beyond scientific risk assessments, and neglects the views on the consumer as economically inefficient due to lack of knowledge and irrationality (Beck, 1986; Bradbury, 1989; Klinke & Renn, 2002). Additionally, the analysis of societal context and determinants is crucial for the understanding of risk amplification and to avoid severe public panics and trust losses as experienced in the BSE crises in Europe in the 1990s (Kasperson et al., 1988; Leiss & Powell, 2004).

Nevertheless, the inclusion of societal concerns in policy remains often "uncomfortable knowledge which is de facto removed from the policy discourse" (Saltelli & Giampietro, 2017). It is not an easy task to integrate relevant criteria to consumers into the technical discourse. While scientific risk assessments use mainly standardized criteria in technical evaluations of health and environmental effects, consumers determine food safety in less tangible ways (Hansen et al., 2003; König et al., 2010). In the examination of risk perception, concerns or safety evaluations, the question arises which kind of measurements and determinants should be used: overestimation and the implied overregulation might lead to welfare losses, for example in forms of higher food prices, in the same way that underestimation or ignorance could (Fraiberg & Trebilcock, 1998; Henson & Traill, 1993).

In 50 years of risk perception research, consumers' safety evaluations and risk perception have been shown to be multidimensional and influenced by various factors. Starting from the findings of the multidimensionality of risk perception, scholars explored the role of psychological processes, culture, socio-demographics, value concepts, and identity building (Douglas & Wildavsky, 1983; Hansen et al., 2003; Kahan, 2016a; Loewenstein et al., 2001; Slovic, 1987). Nevertheless, Wilson et al. (2019) found that many studies which operationalize risk perception remain rather generic. Additionally, closed survey questions are likely to overestimate consumer concerns for food risks (Gaskell et al., 2017). An alternative approach is to look into qualitative criteria which are used in safety evaluations. In risk perception research, a large number of different constructs and terms on different levels can be understood indirectly as food safety criteria (Hassauer & Roosen, 2020). For example, Slovic (1987) has shown dread and familiarity of a hazard

as influential on risk perception, so it is likely that both are used as criteria for evaluating safety. However, only a few studies took a step back and looked inductively and directly at the underlying criteria that consumers actually use for the evaluation of food safety. One example can be found in studies on concerns related to specific hazards (Miles & Frewer, 2001).

The objective of this study is to identify and classify the criteria that consumers use to assess and determine food safety in an inductive approach. We aim to broaden the current perspective by identifying criteria which are used to determine food safety beyond limited concepts such as concerns or risk. For this purpose, we analyze 53 in-depth interviews using a grounded-theory approach. The preceding arguments underline the suitability of a qualitative approach.

The analysis of food safety criteria offers a holistic view on the basis of consumer judgements. The understanding of these judgements is an important task because consumers seem to demand and gain more influence in food-safety discourses. The holistic view opens room for new possibilities of evidence-based inclusion of public food-safety determination in policy through adapted measurements, problem structuring, and framing of food-safety assessments.

3.2.3. Methodology

3.2.3.1. Data

We used a grounded theory approach to identify consumers' food safety determination criteria. The analysis is based on secondary data, originally collected in 2011 in a food-risk study that examined gender differences in hazard-associated meanings. A sample of consumers was interviewed using the laddering technique. These specific in-depth interview form is used to explore relevant product attributes with the linked, underlying consequences and value concepts. Therefore, the interviews cover a wide range of potential food safety criteria and provide deep insights in food-safety determinations. That makes the data a valuable source for this research. In the interviews, participants discussed the risks of herb mixes regarding four differently labelled products. Labels indicated the potential presence of the hazards mycotoxins, pesticide residues, and the use of irradiation (and containing none of them as reference product). The products were

chosen as case studies because they represent the three potential sources of food risks recognized by consumers: microbiological, chemical and technological hazards. The three different hazards enable us to get a holistic view on how consumers determine food safety beyond individual cases (Yeung and Morris, 2001). A more detailed description of the research instrument can be found in the original study (Bieberstein & Roosen, 2015).

The characteristics of the sample are described in Table 11. The participants were recruited by a market research company in Germany, aiming for a representative sample of the German population with quotas on age, income, education and employment status.

Table 11: Sample characteristics (n = 53)

Variable	**Frequency**	**Percent**	**Variable**	**Frequency**	**Percent**
Gender			*Income**		
Male	26	49 %	Low	12	23 %
Female	27	51 %	Medium	26	49 %
			High	15	36 %
Age group			*Food purchasing respons.*		
20 - 35	13	25 %	Main responsibility	37	70 %
36 - 55	21	40 %	Shared responsibility	6	11 %
56 - 80	19	36 %	Not responsible	10	19 %
Education			*Employment*		
Secondary general school	9	17 %	Student	6	11 %
Intermediate secondary school	21	40 %	Homemaker	7	13 %
(Specialized) grammar school	11	21 %	Retired	11	21 %
University	10	19 %	Unemployed	2	4 %
			Employed (full time)	14	26 %
			Employed (part-time)	10	19 %
			Other	1	2%

*Classification: Net monthly household income: Low: <€1200; medium: €1201–3000; high: >€3001.

The broad representation of relevant socio-demographics including gender, age and income increases the likelihood of theoretical saturation (Patton, 2015). According to the purpose of the original study, the comparison of women and men, gender was distributed similarly in the socio-demographic profile.

3.2.3.2. Analysis

For the purpose of this study, we analyzed 53 in-depth laddering interviews using the software MAXQDA. Recorded data and full transcripts were available for the analysis. We used an inductive two-step approach to develop a coding system by applying the Gioia-methodology (Gioia et al., 2012). This approach included a data structuring by identifying criteria close to the text and abstracting them into theoretical themes (equivalent with the coding system) which can be summarized at the aggregate dimensions of food safety criteria. For example, an interviewee talks about the importance of being able to care for his/her children. The first step would mean a first coding as "taking care for children" and the related more abstract, theoretical theme would be "responsibility for family and friends" (for the theoretical themes see also section 3.2.4, results). This methodology was chosen because it provides a systematic and transparent approach for inductive, qualitative research, and hence, improves its credibility and transparency (Gioia et al., 2012). Four research assistants participated in the development of the coding system and coded the interviews after they received specific coding instructions. High intercoder-agreement values in the coding of one interview verify the quality of the developed coding system and the results (code existence and frequency >83 %, segment agreement > 80 %). To further improve the quality, all coded segments were checked for mistakes after final coding by the directing researcher and, if necessary, documented and corrected (removed or re-coded).

3.2.4. Results

In the analysis, we identified five different types of dimensions across the three individual cases. They can further be sorted in either safety-dimension criteria or safety-evaluation criteria (see Figure 8). On the one hand, safety-dimension criteria refer to criteria which are used to express how to frame safety and what constitutes safe food. They can be sub-divided into (I) self-regard and (II) universal criteria. On the other hand, safety evaluation criteria refer to criteria which are related to the evaluation of the safety of food and can be sub-divided into (III) heuristically derived criteria which are used in the evaluation of safety, (IV) expressed responsibilities for food safety and (V) criteria which are stated to have an influence on the evaluation itself.

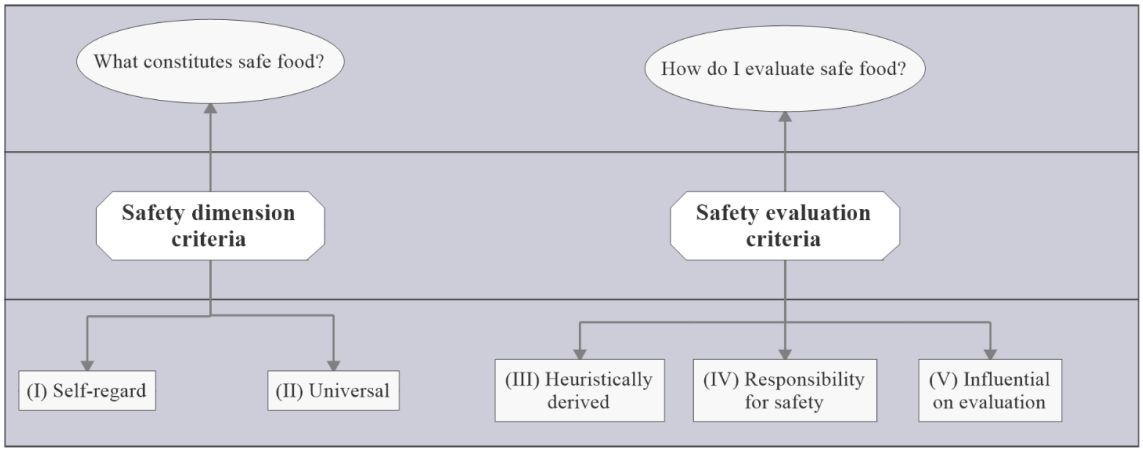


Figure 8: Classification of food-safety criteria

A detailed list of the criteria is provided in Table 12. In Table 12, all criteria are listed, independently from the frequency of them being mentioned. Criteria in (I) self-regard dimensions are used to describe what consumers state as important for safe food and constitute criteria related to themselves. As illustrated in Table 12, safety does not only relate to the negative impacts on physical health but also on life satisfaction, finances, and family and friends. Criteria in (II) universal dimensions relate to mentioned safety criteria which are not related to the individual but to the society and environment as a whole. This includes mainly altruistic concerns for the environment, for other living beings, and for vulnerable groups.

Table 12: Detailed descriptions of food safety criteria

Theoretical dimensions	***Identified Criteria***	***Sub-criteria (optional)***
(I) Safety dimension: Self-regard	Impact on satisfaction and fulfillment	• Healthy lifestyle • Enjoyment of life • Enjoyment of food • Long life/Self preservation
	Financial loss	
	Impact on ability to perform	
	Impact on independence and self direction	
	Impact on health	• Avoid harm • Psychological health effects • Physiological health effects
	Time loss	
	Impact on own physical appearance	
	Impact on being able to be responsible for family and friends	
(II) Safety dimension: Universalism	Concern for other living beings	
	Economic burden on society	
	Concern for vulnerable groups	
	Concern for future generations	
	Concern for environment/earth	
(III) Safety evaluation: Heuristically derived	Connotations	• Positive • Negative • Neutral
	Comparisons to	• Familiar reference points • Other Hazard (case-studies) • Natural status of food • Past/nostalgia • Biographical information
	Indicators	• Seasonality • Freshness • Price • Quality • Country of origin • Local • Organic • Taste/smell • Labeling • Appearance
(IV) Safety evaluation: Responsibility for safety	Self-responsibility	
	Food producers	
	Science	
	Government	
	Food control	
	Independent third parties	
(V) Safety evaluation: Influential on evaluation	External	• Environmental exposure • Dose • Food changes • Benefits • Media

		• Lack of information provided
	Internal	• Sensoric self-evaluation • Knowledge • Lack of knowledge • Lack of understanding

In the safety evaluation consumers use a variety of (III) heuristically derived criteria. In the code book, heuristically derived criteria are defined as determined in any method or process applied by the interviewee to solve a problem, in this case to assess the hazard/technology s/he is interviewed about. This dimension includes comparisons, indicators, and connotations. For example, comparisons can be found to familiar reference points which are often case-study specific. For instance, irradiation is often compared to nuclear power in evaluations, mycotoxins are compared to house mold. Indicators include different product attributes which are stated to be used to evaluate safety, for example the price or taste of a product. Additionally, connotations are used as heuristics for safety evaluations and are mostly negative. In the code book, connotations are defined as following: Following the definition of the Oxford Dictionary, a connotation is "an idea or feeling which a word invokes for a person in addition to its literal or primary meaning". Connotation thus refers to secondary meanings to something suggested or implied by a word or thing (e.g. the words "irradiation" or "mold"). For the fourth type of criteria, consumers use (IV) references to responsibilities for safety. Actors in the food safety determination process are used as reference in the safety evaluation. For example, interviewees can evaluate food as safe with references to the perceived well performance of the food control in Germany. The fifth group of criteria are (V) criteria which are stated to be influential on the safety determination itself. These can be internal, for example the interviewees mention their inability to evaluate safety because of their lack of knowledge. External criteria which influence evaluation are, for example, a reference to the dose of the hazard or the provision of information.

3.2.5. Discussion

A complex set of criteria can be found in consumers' food safety evaluations. Most important is the distinction between criteria which describe relevant dimensions of safety and criteria which are used in the evaluation of safety. The results confirm the role of heuristics, or value-driven determinations, known phenomena from risk perception research, as food safety criteria.

The study has different limitations which need to be considered when interpreting the results. We use secondary data that is very suitable for the objectives of the study. Nevertheless, the underlying research design, oriented at the means-end-chain theory, focusses on a certain type of criteria (attributes, consequences and values) and might exclude other types of criteria which are used besides these concepts when evaluating and judging food safety. Additionally, the discussed products, dried Italian herbs, might have had an influence on which criteria were mentioned (e.g. freshness or country of origin as indicators). The criteria are listed independently of their frequency of being mentioned, so some of them might be rather irrelevant. A further analysis of code frequencies or a quantitative follow-up study could give insights to the relevance of the elicited concepts.

The study offers a new approach to analyze consumers' safety evaluations inductively. The rich variety of criteria confirms the need for the understanding of qualitative dimensions of safety (Leiss & Powell, 2004). It offers a possibility to provide insights in a "socially robust universe of possible frames, which represent different lenses through which to perceive what the problem is" (Saltelli & Giampietro, 2017). The complexity of the identified criteria confirms that it may be insufficient to measure consumer safety determinations in unidimensional risk perception scales if the objective is to evaluate consumer determination in complex food safety issues (Wilson et al., 2019). It turns out to be important to differentiate between the safety dimension criteria and the safety evaluation criteria.

On the one hand, safety dimension criteria can be useful to determine framings of safety, setting priorities or "determine the criteria of judging tolerable levels of risk, whereby the technical assessments are used as one important input among others to quantify the extent of potential damage in time and space" (Klinke & Renn, 2002). They can be linked to risk perception studies studying underlying values or specific concerns and confirm their role as type of food safety criteria inductively (Bieberstein & Roosen, 2015; Miles & Frewer, 2001; Slimak & Dietz, 2006). The food safety evaluation criteria, on the other hand, are used for evaluations of safety and can give hints on potential biasing elements which can lead to under- or overestimations of food risks by the public. A lot of them are already described in literature as influential on risk estimations (e.g. Lusk et al., 2014; Sjoberg, 2000; Sunstein, 2002). Interestingly, consumers seem to be well aware of these heuristics and state them in the in-depth interviews, for example the use of indicators or

comparisons with known phenomena or an ideal past. This indicates that the analysis of used heuristics might be helpful to optimize communication, interpret safety evaluations or risk estimates. Additionally, the references to responsibilities for food safety can be useful for the interpretation of trust scales or the optimization of risk communication strategies (Erdem et al., 2012; Redmond & Griffith, 2004; Sapp et al., 2009)

3.2.6. Conclusion

Our findings indicate that it can be useful to identify underlying food safety criteria inductively rather than to measure aggregated risk perception or concern scales in complex or contested food safety discourses. Aggregated measurements can neither account for the complexity of food safety determinations nor identify the specifically used criteria. A distinction between the safety dimension criteria and the safety evaluation criteria can enable regulators and researchers to identify criteria which are used in the framing of safety (dimensions) and heuristics and potentially biasing elements in the evaluation of safety (evaluation). This can be useful for the integration, interpretation, and communication. Additionally, differences in risk perception between hazards can be expressed in multidimensional criteria which acknowledge the underlying sources for these determinations. A further analysis of the data can provide insights into relevance of criteria or the identification of hazard-specific criteria for the evaluation of different technologies.

3.2.7. Acknowledgements

This article is based on research supported by funding from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Research Group 2448. We thank Andrea Bieberstein for the provision of the valuable data. We further thank the research assistants Tina Cleffmann, Magdalena Eichhorn, Julia Hörmann and Jovita Jahn for their work in the development of the coding system and the analysis of the data.

3.3. Stakeholder Integration in the Determination of Evidence for Food Safety—Insights from European Risk Assessment Organizations

3.3.1. Abstract

Determining relevant evidence is an important aspect in food safety regulations. Risk assessment organizations play a key role in this process. Although the need for stakeholder integration in evidence determination is widely acknowledged, the participation of stakeholders in these organizations remains a controversial topic because it might, for example, affect the independence. The objective of this study is to identify whether and how members of risk assessment organizations interact with stakeholder groups in selecting and using evidence and thus contribute to the in- or exclusion of information. We provide results from interviews with members from two official organizations in Europe in the fields of plant protection products and microbiological hazards. The qualitative approach allows us to identify different one-way and two-way interactions with two stakeholder groups. Interactions appear on different levels in the selection, processing, and use of evidence. We describe these interactions from the perspective of risk assessment organization members, including practical difficulties and evaluations. The study highlights the importance of identifying the different roles of stakeholders in the selection and use of evidence. Our insights show that it is helpful to put more emphasis on existing efforts and to provide a clear guideline for interaction processes at the different stages of evidence processing and use. This might help to increase stability and trust as well as to consider ambiguity in risk assessments. It is also helpful for supporting risk assessors in the resolution of value conflicts and practical difficulties which are linked to stakeholder participation.

3.3.2. Introduction

Risk assessment organizations are a central element in defining food safety regulations. They are responsible for the provision of relevant evidence to decisions makers[26]. However, judging "relevant evidence" is complicated. Different stakeholders are participating in the food system (Ingram, 2011). Stakeholders are affected differently by food safety: consumers can lose healthy life years in significant numbers or food producers money for compensation or prevention measures (Henson & Loader, 2001; Hessing et al., 2015; Hussain & Dawson, 2013; WHO, 2015). Additionally, stakeholders bear different responsibilities for the management of risks in food chains, including for example through the production, processing, distribution, marketing, and preparation of food (Trumbull, 2006; van der Meulen & van der Velde, 2004). They are able to generate information regarding food safety themselves, for example when monitoring production or voicing complaints. Food safety can be defined as the absence of hazardous substances (e.g. Herges et al., 2017), which is an rather unrealistic scenario. Thus, in the context of this paper, it describes an acceptably low certainty or probability of adverse effects (Hassauer & Roosen, 2020; OECD, 1993). Inevitably, stakeholders need to be integrated in the process of evidence generation for acceptable food safety levels, whether as group at risk, object of observation, or as provider of information.

Despite this apparent importance, the role and extend of stakeholder groups in risk assessment organizations has not been treated in much detail so far. We examine the current inclusion processes for information from two stakeholder groups in the practice of selecting and using evidence in two risk assessment organizations. This analysis will help to understand how stakeholder integration already works, systematically identify current practices, and point to potentially problematic elements. Additionally, it will help to understand the risk assessors' perspective and to highlight their role and difficulties in these discussions.

To develop such a theoretical description, specific research questions are:

[26] In the context of this paper, we understand relevant as "having significant and demonstrable bearing on the matter at hand" (Merriam-Webster, 2021). Evidence is understood as decision-relevant knowledge, not only based on data but more on socially accepted knowledge (Cartwright, 2006; Hassauer & Roosen, 2020; Kelly, 2016).

- How do risk assessment organizations interact with stakeholder groups when selecting and using relevant evidence for food safety?
- Which criteria are used to decide on the inclusion or exclusion of information from stakeholder groups?
- How do experts deal with related issues?

The study does not aim to depict the whole process of evidence determination for food safety in risk assessment organizations but focusses on stakeholder interactions in this process. It covers risk assessment, stakeholder engagement, and risk communication activities in these organizations.

3.3.3. Background

3.3.3.1. Risk assessment as central element in regulating food safety

To understand the relevance of stakeholder integration, we provide some basic information on the role of risk assessment organizations in evidence determination for food safety. In the European Union (EU), relevant evidence for food safety is determined within the risk analysis process which consists of three elements: risk assessment, risk management, and risk communication. All elements fulfill different roles but do need to interact (Devos et al., 2019)[27]. Risk analysis in the EU is characterized by an institutional separation of scientific risk assessment and political risk management, the European Food Safety Authority (EFSA) supported by national authorities being charged with the risk assessment and the European Commission, member state authorities and the European parliament signing for risk management. Risk managers are responsible for problem formulation for risk assessment and decision-making, for example prioritization of microbial hazards or the approval of plant protection products. Decision-making is based on risk assessment but also on societal, technical, and economical contextual information (EC & EP, 2002; Millstone, 2009; NRC, 2009).

[27] Basic risk analysis principles and processes are designed, conducted, and harmonized based on international guidelines and standards, for example the Codex Alimentarius or the International Plant Protection Convention (WTO, 2021b, 2021a).

Risk assessment functions as a means for providing scientific evidence for decision making. The National Research Council describes risk assessment as "the use of the factual base to define the health effects of exposure of individuals or populations to hazardous materials and situations" (NRC, 1983). Any risk assessment consists of four standardized steps, hazard identification, hazard characterization, exposure assessment, and risk characterization (Barlow et al., 2015). In literature, risk assessment is described as a "rational and objective picture [of] what is known or believed to be known" (FAO & WHO, 2009), a "knowledge-generating activity" (Vareman & Persson, 2010) providing "objective evaluations of all available information" (Wentholt et al., 2009). Although not responsible for decisions (Devos et al., 2019), it functions as a decision support tool but not as a scientific method. It is useful for identifying research gaps (FAO & WHO, 2009), organizing data, and allocating responsibility for the analysis (Consultation, 1997), and is seen as a way to lend credibility to decisions (Devos et al., 2019). In these definitions, the assumed level of independence of risk assessment organizations varies. Nevertheless, all of the preceding descriptions entail problematic assumptions: First, while the assumption that a risk assessment is more or less free of societal influences seems rather idealistic, scientific risk assessment is often presented as an independent factor in decision making and a way to represent an ideal picture of scientific values (König, 2007; Millstone, 2009; Vareman & Persson, 2010). Second, this description neglects the existence of different types of ambiguity in risk assessments which include the existence of "multiple interpretations (…) of risk information" (Johansen & Rausand, 2015). Third, it is unrealistic to assume that risk assessments are able to provide all available knowledge about complex risk issues (FAO & WHO, 2009; Fischhoff, 1994).

Risk communication is responsible for the dialogue between risk assessors, risk managers, and external actors (Devos et al., 2019). Thus, risk communication is an important field in risk assessment organizations: For example, EFSA has a separate department which is responsible for communication and engagement (EFSA, 2020).

3.3.3.2. The need for including stakeholder evidence in risk assessments

Most food hazards have multiple effects. Chemicals are produced because they have benefits for some stakeholders but increase risks for others. Risks may entail different trade-offs for different individuals or groups (Fischhoff, 1994; Fischhoff & Hope, 1984).

Other hazards, for example hazardous bacteria, might be undesired but resources to eliminate them are limited (Ruzante et al., 2010). Determining acceptable risk and thus safety levels influences the outcomes of policy, the allocation and prioritization of resources, and regulations (Fischhoff, 1994; Fischhoff & Hope, 1984; Klinke & Renn, 2002). Those affect food producers and consumers in different ways, for example through implications for production techniques, taxes, trade restrictions, food prices, or protection of health. Therefore, acceptable levels of risk in food safety are determined differently by different actors, depending on the consequences which are in- or externalized (Hassauer & Roosen, 2020; Henson & Traill, 1993). It is not a trivial question whether and how information on one group is excluded because the exclusion of it does also exclude the group and its needs or beliefs from resource discussions. Although these decisions of in/exclusion are in the responsibility of the risk management, risk assessment is responsible to provide the relevant information for these decisions. Risk assessment is furthermore affected because these trade-offs contribute to the existence of ambiguity in risk assessments (Johansen & Rausand, 2015; Stirling, 2008). Additionally, inadequate framing of evidence determination by policy can problematize risk assessments and force risk assessors into value judgements (Devos et al., 2019; Vareman & Persson, 2010). EU Principles of Good Governance require a broad assessment including economic and social effects, but reality is different. Generally, responsibilities and processes for the inclusion of social or economic information are defined insufficiently in current risk analysis frameworks including risk assessment organizations (Barker et al., 2010; EFSA Panel on BIOHAZ, 2012; König et al., 2010; Ruzante et al., 2010)[28].

One way forward is to integrate this type of information via stakeholder participation in risk assessments. Stakeholder involvement in risk analysis ensures satisfactory decision processes and are relevant in all stages of the decision-making process (NRC, 2009). Stakeholder participation can appear in different forms, and can have different target groups or objectives, but the extent of its current use is not easily observable. It appears

[28] Although this criticism was raised a decade ago, there has been no significant change toward systematic integration of stakeholders or social acceptance of risks in regulatory practice so far (Devos et al., 2019; Morvillo, 2020). A positive example is the integration of stakeholder participation in the EFSA 2020 strategy. Nevertheless, its major objectives are the establishment of trust, engagement and dialogue but not necessarily the concrete integration of stakeholder evidence in the risk assessment process (EFSA, 2016a).

as involvement, participation, or engagement. Most literature focuses on public participation as part of the risk management and risk communication processes (e.g., see Rowe and Frewer (2005) for a categorization of public participation mechanisms). However, in- and exclusion of stakeholders is also relevant for risk assessment through the framing of the endpoints of risk assessments and because risks are influenced by actors through their production and consumption decisions. Therefore, risk assessment also needs to include stakeholder information on these actions. It is acknowledged that public discourse is likely to influence systematic knowledge production (Bayley & French, 2008; Renn, 2015; Vareman & Persson, 2010). To the best of our knowledge, there is no systematic analysis of stakeholder interaction in risk assessment for food safety. That is problematic because it complicates identifying, developing, and effectively applying and regulating such activities (König, 2007; NRC, 2009; Rowe & Frewer, 2005).

Preliminary work on the systematic categorization of stakeholder interaction is undertaken by Rowe & Frewer (2005). They classified public engagement mechanisms based on the flow of information between public representatives and the commissioning party (sponsor). These engagement mechanisms are defined as public communication (one-way information flow from sponsor to public), public consultation (one-way information flow from public to sponsor), and public participation (two-way information flow between public and sponsor) (Rowe & Frewer, 2005). We base our analysis on that same logic and aim to identify these types of interactions in the activities of risk assessment organizations in food safety.

Risk assessments deal with complex topics, uncertainty, quantitative probabilities, and controversies. Consumer knowledge, for example, is different in quality and difficult to include in a traditional expert culture (Barker et al., 2010; Fischhoff & Hope, 1984; Homan et al., 2001). That entails practical difficulties for risk assessors that need to be resolved and clarified in stakeholder participation. Therefore, one objective of this study is to describe the related context and difficulties for risk assessors.

The following section describes the case study research strategy including the choice of cases and the description of data collection and analysis.

3.3.4. Material and methods

3.3.4.1. Research strategy

An embedded, multiple case study was chosen as research design. Case studies are suitable for understanding real-world cases, including the assumption that this understanding involves contextual considerations and that boundaries between case and context are not clear (Yin, 2013). Units of analysis are interactions in the integration of (evidence on) (1) consumers, and (2) the food sector in food risk assessment organizations in a European context. We use interviews with members of risk assessment organizations to observe how these interactions work in actual practice. We chose one national (German) and one supranational (European) risk assessment organization. Both organizations work under the same regulatory framework, the General Food Law, but fulfill different tasks (EC & EP, 2002).

We selected two groups of stakeholders relevant in the process of food safety determination to assure a broad perspective on the topic: supply side (food sector[29]) and demand side (consumers[30]). Both stakeholder groups are important actors within the food system (Ingram, 2011). They have different possibilities for influencing risks: Consumers are seen as able to inform themselves and handle food with care (van der Meulen & van der Velde, 2004), face an information asymmetry because they have few insights into food production processes (Trumbull, 2006), and have a different practice for determining evidence that is difficult to bring in line with scientific food safety determination (Hassauer & Roosen, 2020). The food sector is responsible for complying with regulations throughout all stages of production for all food products (Vos, 2000).

We used a maximum variation sampling strategy (Yin, 2013) by selecting the cases of microbiological hazards and plant protection product residues. We did so for various reasons: First, both hazards are different in nature. While it is possible to decide how to introduce and use man-made plant protection products, microbiological hazards occur naturally and behave highly dynamically (IoM & NRC, 2003). Second, consumers

[29] Here, the food sector as a stakeholder group is seen as a broad term that includes primary production, food processing, retail, and certification.

[30] Here, the group of consumer stakeholders includes individuals and consumer organizations.

perceive both hazards very differently. While the risk from chemical hazards is generally overestimated, microbiological risks tend to be underestimated (Yeung & Morris, 2001). Additionally, safety determination in both cases are stipulated in different European regulations but have a long history of risk analysis (van der Meulen & van der Velde, 2004). The previous arguments support the expectation that both cases can contribute different aspects regarding interaction procedures and dynamics.

3.3.4.2. Data collection

We conducted interviews with members of risk assessment organizations. These members were selected because they have unique insights into this topic and the related processes. Therefore, they are defined as experts due to their technical, process and interpretative knowledge (Bogner et al., 2009)[31]. Selection criteria for experts were the institutional affiliation (official national and supranational risk assessment organizations in the European food safety framework) and their work in a case-study related field (plant protection products and microbiological hazards) plus risk communication. Risk communication was included because it is an important element for collecting evidence on consumers in risk assessment organizations, for example, the provision of consumer insights into the Eurobarometer surveys (EFSA, 2019). Risk assessment and risk communication work closely together, i.e., in the same organization, and exchange information and opinions in an interactive process (Devos et al., 2019; Wentholt et al., 2009). Therefore, it can be assumed that risk communication provides a certain amount of consumer-related evidence related to risk assessment. Interview partners were partly addressed based on publicly available organigrams, and partly presented by the institution after official inquiry and information on the interview content. In total, the sample consisted of ten members of risk assessment organizations in different fields of expertise (s. Table 13).

[31] We are aware that experts cannot be seen as neutral sources of information independent from their tasks in risk assessment institutions. This point should be kept in mind in the interpretation of results and is elaborated in the discussion section.

Table 13: Overview of interviewees

ID	Field of expertise/ Working area	Years of experience in (similar) position at point of interview	Gender
1	Risk communication	11	M
2	Plant protection products	11	F
3	Plant protection products	19	M
4	Microbiological hazards	+/- 5	M
5	Microbiological hazards	<1	M
6	Stakeholder engagement	3	M
7	Risk communication	9	M
8	Plant protection products	14	F
9	Microbiological hazards	5-6	M
10	Microbiological hazards	12	M

We used open and semi-structured, guide-based interviews to create a more natural conversation setting and to allow the focus to depend on the different forms of expertise. Interview guide approaches are broadly applicable as long as the aim is not to quantify statements (Gläser & Laudel, 2010). The interview-guide approach has the advantage of producing comprehensive data, enabling systematic data collection, but simultaneously allows interviewers to remain situational and flexible (Meuser & Nagel, 1991).

The content of the interviews is based on the research questions above and focuses on interactions between risk assessment organizations and stakeholders (consumers/food sector) (see Figure 9).

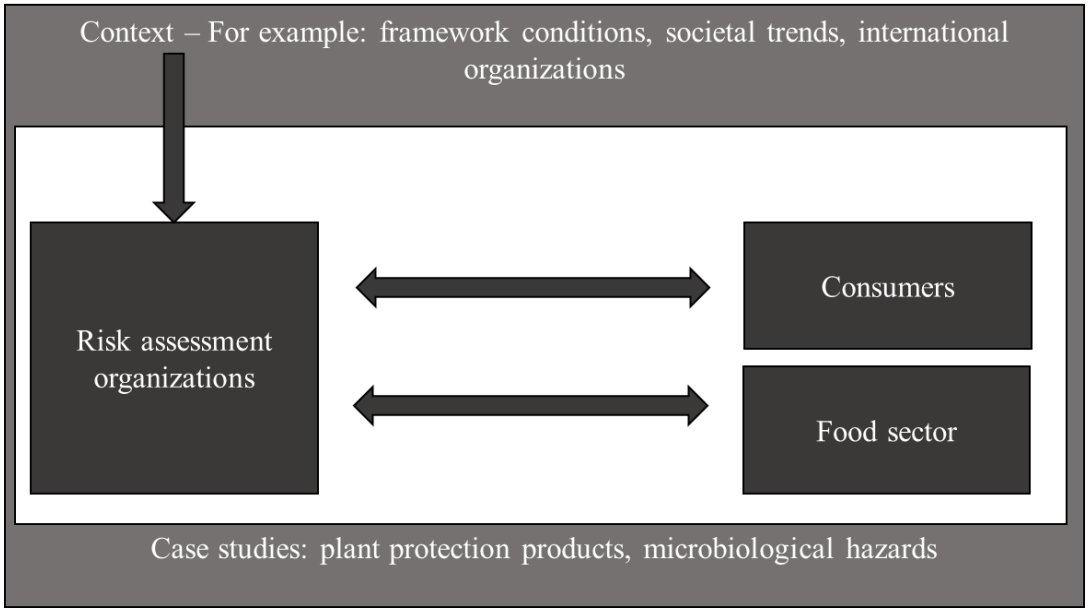


Figure 9: Content structure interview guide

Interactions are characterized by uni- and bidirectional exchanges of information (Rowe & Frewer, 2005). Therefore, the first set of questions targeted the identification and potential integration of (1) collected information on stakeholders by risk assessors, and (2) information produced by stakeholders themselves. It is assumed that this information is diverse. Sub-questions targeted identifying quality criteria to in- or exclude information as "stabilizers" during the information selection processes. Including stakeholder information can be challenging (Barker et al., 2010; Fischhoff & Hope, 1984; Homan et al., 2001). Sub-questions focused on potential difficulties in using this information that can turn into potential "destabilizers" of the risk assessment processes and obstacles for the work of risk assessors and communicators. Questions were designed based on the Patton-Matrix considering the time and type of question (Patton, 2015). The interview guide was slightly adapted for the interviews with risk communicators. Both versions can be found in Appendix C.

The interview guides with most of the guiding questions were sent in advance to allow the interviewees to prepare. Some questions were posed during the interviews exclusively because we aimed for spontaneous reactions.

3.3.4.3. Data analysis

The interviews were audio recorded, fully transcribed, and sent to interviewees for release if requested. The transcripts were anonymized and transferred to the software *MAX QDA*.

Every qualitative analysis is unique so we mixed two analytic approaches according to our purposes (see Figure 10) (Patton, 2015). The analysis is based on a specific approach for expert interviews by Meuser & Nagel (1991, 2009) and specified through a coding technique by Gioia et al. (2012). Both approaches use the principle of two coding rounds and a round of abstraction, so they complement each other very well.

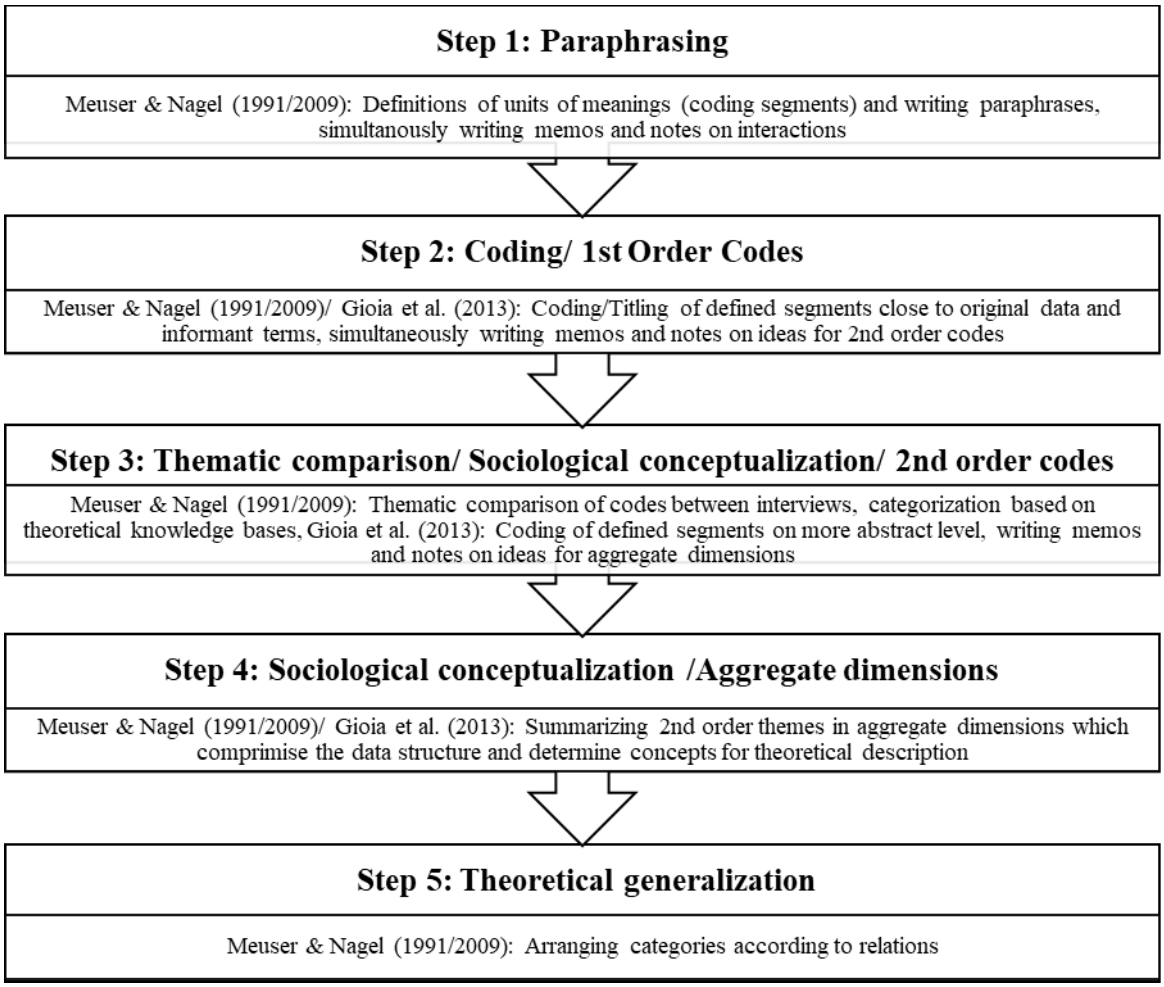


Figure 10: Structure of analysis (Source: Gioia et al., 2012; Meuser & Nagel, 1991, 2009)

The combination of these two approaches has various advantages. The small steps of the Meuser & Nagel (1991, 2009) approach and the extensive description of coding techniques of Gioia et al. (2012) increase transparency and reliability, which is a major issue, especially in semi-structured in-depth interviews[32] where often only one coder is

[32] Common problems in the process are: Not much guidance for process, problems of availability of multiple coders: need for knowledgeable coders, need for background knowledge (Campbell et al., 2013)

available. Additionally, it improves the recursiveness of the analytic process (Meuser & Nagel, 2009).

We defined units of meanings as coding segments (Campbell et al., 2013), and we did not cut off codes with small counts when we presented our results because it is the aim of this expert study to assemble the available knowledge and all its aspects (Meuser & Nagel, 1991).

As a result, we identified themes and sub-themes, for example, the type of collected data on stakeholders. We structured the results in an own theoretical description of the cases. An overview is presented in Figure 11 in the following section. The theoretical description is a result of theoretical generalization (Meuser & Nagel, 2009). It shows cross-case patterns (not cross-case comparisons) and provides a joint description of both cases, microbiological hazards and plant protection products as well as interactions between risk assessors and both stakeholder groups (Yin, 2013). Themes and sub-themes are presented in Appendix D. We are aware that sub-themes might be incomplete. Hence, they rather serve as examples for main themes and might be extended with different cases.

3.3.5. Results

The aim of this study was to identify and describe interactions of risk assessment organizations with two different stakeholder groups (food sector/consumers) in selecting and using evidence. Figure 11 provides insights in two dimensions of interest. Vertical arrows show interactions with the stakeholder group of interest in selecting and using evidence. Horizontal arrows display the use of information on stakeholders within the risk-assessment institution.

The theoretical description in Figure 11 will be elaborated and explained in the three following sections following the horizontal logic from left to right: information input, processing of information, evidence output.

Horizontally, the graph starts on the left with input in the process, more specifically with stakeholder-related information used in risk assessment organizations. The center displays processing of information in (a) risk assessment, and (b) risk communication,

more specifically defining relevant variables used in models to determine risk metrics[33]. In this section, we present risk assessment and communication separately because interviewees describe an information exchange between the two fields, which is elaborated in the following sections. On the right, we describe evidence outputs generated in risk assessment and communication, for example, in form of scientific opinions. We provide detailed descriptions and examples in the following sections.

[33]The terms are defined based on the structure described in EFSA Panel on BIOHAZ (2012). Here, variables are defined as "factors which should be considered in decision-making regarding an issue."

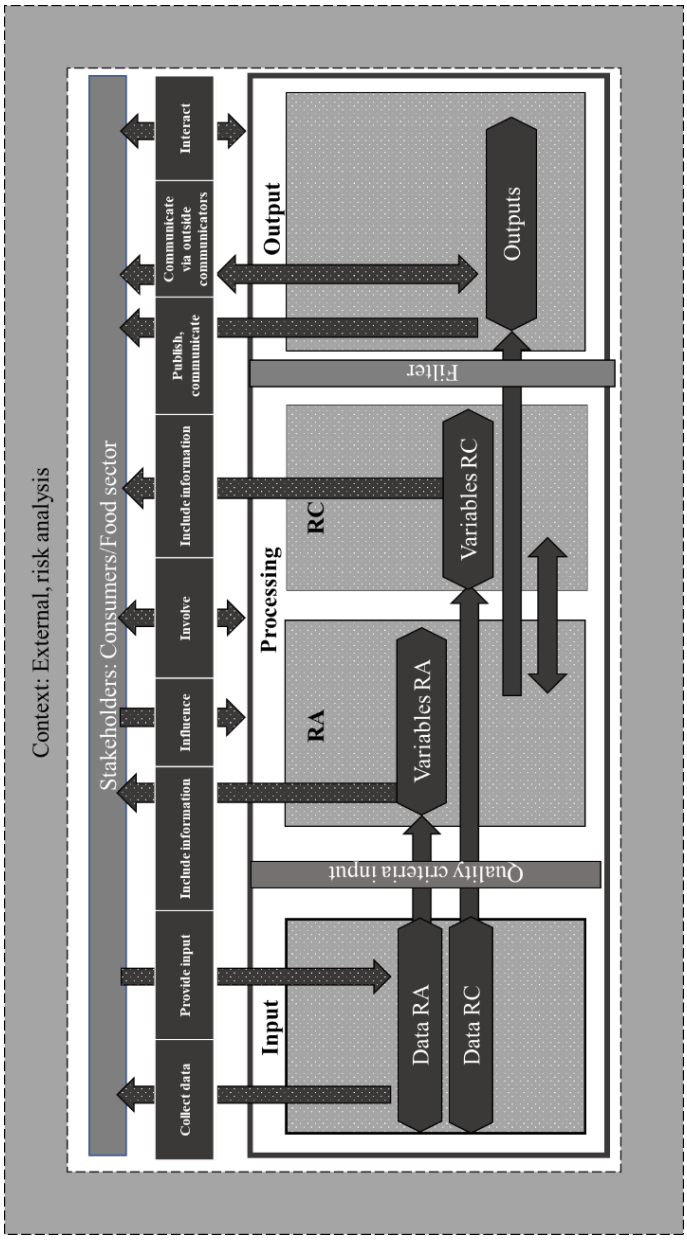


Figure 11: Theoretical description of the cases. RA = risk assessment, RC = risk communication

3.3.5.1. Information input

The first element is the input in the risk assessment/communication process, more specifically, the inclusion of stakeholder-related data or information to determine relevant effects.

Interactions with stakeholders

We start with vertical arrows in Figure 11: They describe input-related interactions with stakeholders. The arrows represent the direction of force for the initiative rather than the

information flow per se, which always flows from stakeholders to the inner box (risk assessment and communication). We identified two types of interaction—the collection of data by the institute and the provision of input by stakeholders. A detailed list of all sub-themes can be found in Appendix D. Figure 12 presents the overview of the direction of force for initiative.

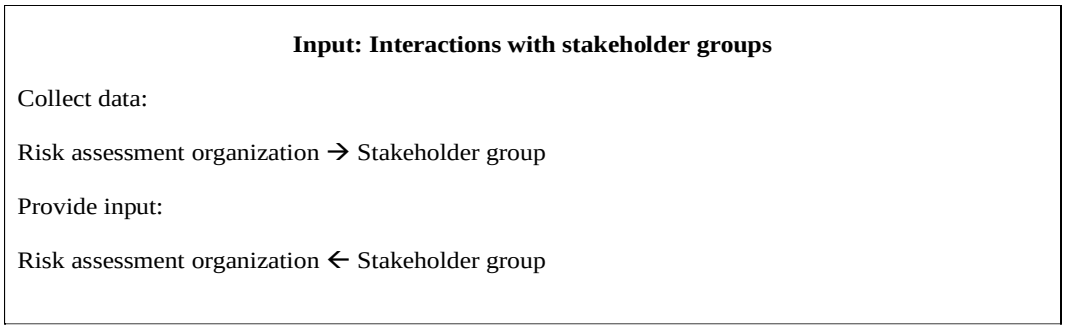


Figure 12: Input: Interaction with stakeholder groups

First, "collect data" summarizes types of data collections initiated by risk assessment organizations on stakeholders, for example, focus groups (sub-theme: data collected on consumers by risk communication) or monitoring (sub-theme: data collected on food system by risk assessment). Data collection does not automatically imply that risk assessors collect the data themselves but may also mean that they use existing data or commissioned studies. Stakeholders hold a position as passive objects of interest for the risk assessment organizations because they are consulted or observed with specific questions. Risk assessment uses various, diverse tools to collect data from the food sector, for example, by specific consultations or lab studies on production techniques. For consumers, experts stated the use of surveys, data bases and literature. In risk communication, types of collected data are more diverse and also include qualitative approaches, for example, those based on focus groups. Types of selected information vary not only between stakeholders (food sector or consumers) but also between purposes (risk assessment or risk communication).

Second, the provision of input describes information input provided by stakeholders. In this interaction, stakeholders hold the position as active provider of information to the risk assessment institution. This does not automatically imply that the data covers information on a stakeholder but that the data is provided by a stakeholder. Examples are public consultations (sub-theme: consumer-to-risk assessment) or provision of self-

monitored data (sub-theme: food-sector-to-risk assessment). Provided data leads to evidence of various qualities. Experts described the problem of non-scientific input provided by consumers, which remains irrelevant in the risk assessment process. The food sector generates huge amounts of valuable data but has a limited willingness to share it with risk assessment organizations. Experts explained this unwillingness of conflicting interests of transparency and privacy of information but report an increasing willingness to share data due to growing public expectations.

It became apparent in our analysis that qualitative approaches in consumer research are only used in risk communication. On the food sector side, interviewees critically discussed the provision of self-monitoring data[34]. In general, we find different types of data that are challenging to combine. This is also visible in the variety of quality criteria used, which are discussed in the following section.

Use of information in risk assessment organizations

We continue with the description of the process along the horizontal arrows in Figure 11. They show the use of information in the process of risk assessment and risk communication. Arrows pass a barrier, which we title "quality criteria" or "filter"[35]. Quality criteria are used to ensure input quality and to decide about in- or exclusion of input and the ultimate assurance of quality. They can also include statements related to flaws in data (data quality not sufficient, assessed by criterion). Quality criteria of risk assessment and risk communication differ (for a detailed list please see Appendix D), an interviewee reported that they are stricter in risk assessment. In risk assessment, one can find "classic" scientific quality criteria such as peer-review; context-related quality criteria (sub-themes: data pool dependent on context, consideration of costs of studies), reference to own expertise as quality criterion, and external expectations such as independence and transparency. Risk communication experts reported context-related quality criteria, "classic" scientific quality criteria, and reference to own expertise.

[34] For example, experts critically discussed the partly low willingness to share data (but positive trend), mentioned the danger of economic damage to sharing businesses, or reported the opportunities and problems of the large amounts of available data.

[35] The term "filter" and the logic as barrier is adopted from the systematic use of evidence processing in policy making described by Vogel & Delfini (2008).

Interesting insights are for example the description of a variety of quality criteria which are partly implicit and go beyond classic scientific quality criteria. One example is that including and using data depend on context and the available knowledge base or the relevance of societal expectations such as independence. Also, active input by stakeholders has to pass the quality barrier. The strong reliance on scientific quality criteria might make it difficult for non-scientific information provided by consumers to pass that barrier into the use and processing of evidence, which is described in the following section.

3.3.5.2. Processing of information

Interactions with stakeholders

In this section, we describe interactions with stakeholder groups related to the processing of evidence in risk assessment and risk communication. We start with vertical arrows in Figure 11. Results show three different types of interactions with stakeholders when processing evidence illustrated in Figure 13.

Processing: Interactions with stakeholder groups

Include information:

Risk assessment organization → Stakeholder group

Influence:

Risk assessment organization ← Stakeholder group

Involve:

Risk assessment organization ← → Stakeholder group

Figure 13: Processing: Interactions with stakeholder groups

First, risk assessors and communicators include information on stakeholders in the form of variables[36] that are derived from previously described information input. Stakeholders hold a passive position; they are included through variables that contain relevant

[36] "[F]actors which should be considered in decision-making regarding an issue," (EFSA Panel on BIOHAZ, 2012) which are used in models to determine risk metrics. Categorizing information flow is generally derived based on structures described in (EFSA Panel on BIOHAZ, 2012).

information on them. Results show a variety of variables which are listed in Appendix D. Perception and acceptance do not play a role in risk assessments (some experts reported possibly indirect consideration via behavior) but various other variables on consumers are included. Risk assessment considers, for example, behavior, consumption or knowledge. Food sector-related data includes, for example, production hygiene or product volumes. Experts report that food sector-related data does not play a role in safety determination of regulated products (plant protection products). Risk communication includes a different set of variables, for example, perception, trends or information sources. A possible indirect consideration of these variables in risk assessment through an exchange is described inthe preceding section "*Information exchange between risk assessment and risk communication"*.

Second, stakeholders actively influence the process of risk assessment and risk communication in different ways. These influences are less explicit than the passive consideration in form of variables. Experts report, for example, that consumer requests bind resources or initiate changes in process or method development. Experts describe active influence by food sector as mostly irrelevant in risk assessment, and have different opinions on the amount, relevance and value of consumer influence on the practices in risk assessments. Some interviewees hold positive opinions on these (attempted) influences, whereas others evaluated them as negative[37]. Several instruments ensure that consumers have the option of approaching risk assessment institution, for example, hotlines or dialogue events.

Third, experts reported a two-way interaction, that is, the stakeholder engagement. It involves stakeholders in risk assessment organizations in a two-way communication process, for example, in consulting stakeholder working groups. Establishing trust was reported as a major objective. The stakeholder engagement is considered in risk assessment and communication, but experts do not report a direct influence on the work

[37] For example, opinions of interviewees differed in the following case: A new analytic method was developed after public concern on this specific topic. Interviewees discussed whether the invention of the advanced risk-assessment quality method or whether its development was an unnecessary use of money and resources.

of risk assessment organizations, one expert described the influence as difficult to measure.

As described in this section, variables vary between risk communication and risk assessment. Therefore, it is interesting how evidence is exchanged between both fields. We describe these exchanges in the following section.

Information exchange between risk assessment and risk communication

To read Figure 11, we proceed with the horizontal arrows in the center. They display the interaction between risk assessment and risk communication, which can take various forms. Experts state, for example, translation of information, raising awareness or information exchange. These interactions show that stakeholder evidence from risk communication (for example, on risk perception) plays an indirect role in risk assessment. The variety of interactions demonstrates a close link between risk communication and risk assessment in the exchange of information. On the right side of the processing section, evidence "leaves" the risk assessment institution as output that is described in the next section.

3.3.5.3. Evidence output

Interactions with stakeholders

In this section, we describe interactions with consumers in the production and use of outputs of risk assessment organizations. We start again with vertical arrows in Figure 11. We do not distinguish between risk assessment and risk communication in this section because outputs are generated mostly in cooperation. Results show three different types of interactions with stakeholders with regard to outputs of risk assessment organizations, which are illustrated in Figure 14.

Output: Interactions with stakeholder groups
Publish, communicate:
Risk assessment organization → Stakeholder group
Communicate via outside communicators:
Risk assessment organization → Outside communicator → Stakeholder group
Interact:
Risk assessment organization← → Stakeholder group

Figure 14: Output: Interactions with stakeholder groups

First, we find publishing of outputs and communication to stakeholders. This includes different publication formats addressed at various publics, for example, risk assessments or scientific opinions but also social media content. Included in this theme are issues that arise in communication with the public for example the difficulty to explain thresholds, or biased perception of different hazards. A detailed list of sub-themes is provided in Appendix D.

Second, we find an indirect interaction via outside communicators, namely media and multiplicators (for example, consumer organizations). Most experts reported this interaction as a critical element. It was identified as an issue that risk assessment organizations do not have a large reach and that most of the outputs are spread through outside communicators, multiplicators or media. While multiplicators are seen as "partners" (exchange, direct line), media is seen as important but was more critically discussed as exemplified in sub-themes (for example, sub-themes: fuel public fears, distribute fake news, attack risk assessment organizations, (mis-)interpret information).

Third, we find a two-way interaction with consumers, for example, in the evaluation of communication tools (feedback) and interaction with the public on social media platforms.

The output summarizes the selected and presented evidence of a risk assessment institution and can be presented in various forms. This form is determined by a number of quality criteria or filters, which are elaborated on in the following section.

Use of information in outputs

We continue with the description of horizontal arrows in Figure 11. Evidence crosses a barrier titled "filters"[38]. These include quality criteria that decide which and how information is provided as output to actors outside the risk assessment organizations. We find, for example, contextual conditions for the presentation of outputs (for example sub-themes: form depending on target group or regulatory framework—for a detailed description, please see Appendix D).

3.3.5.4. Context

The last part of the theoretical description of our case studies are themes related to the context of the unit of analysis (interactions of risk assessment organizations with stakeholders) and are displayed as a frame in Figure 11. As defined by Yin (2013), cases cannot be separated from their context. This allows and requires a context-related description of the cases. We described some contextual factors as part of quality criteria and filters in the related sections. Besides those, we find more general contextual conditions, which are presented in the following section. We found several themes in the interviews that are related to contextual conditions for stakeholder interactions in generating and using evidence in risk assessment organizations. We partitioned context-related themes into the risk analysis context and external context.

Experts report external influences from different societal actors. Nearly all sub-themes describe how different actors set impulses for stakeholder action: For example, interest groups criticize and publish own evidence. Science, policy or risk management initiate critical debates on generating, selecting and using evidence.

Influences from the general risk analysis context are mostly related to the initiation of amendments in risk assessment practices or the institutional frameworks, which restrict the ability of risk assessors to react to external criticism. Interviewees report, for example, that risk management is responsible for framing risk assessments and therefore, the reaction to outside criticism.

[38]The term "filter" and the logic as barrier is adopted from the systematic of evidence processing in policy making described by Vogel & Delfini (2008).

3.3.6. Discussion

Risk assessment organizations are an important element of food safety regulation for integrating relevant evidence into a current state of knowledge. Although risk assessments aim to provide independent and objective advice, it is unrealistic to think that society, stakeholders, or value considerations do not influence the practice of risk assessment (Devos et al., 2019). Stakeholder participation is needed and undeniable (Robinson et al., 2016). Nevertheless, the inclusion of stakeholder information is a challenging task in the work of risk assessment organizations. Stakeholder participation is not a clearly defined term and can include various activities at different levels (Rowe & Frewer, 2005). The aim of this study was to identify and to systematize interactions of risk assessment organizations with two defined stakeholder groups (consumers/food sector) in selecting and using evidence for food safety in two different cases. It was possible to identify basic cross-case patterns and to develop a joint theoretical description although we chose a maximum variation sampling strategy. This confirms that we were able to provide a literal replication and might indicate that the basic description of interactions is suitable for other food safety cases as well (Ridder, 2017; Yin, 2013).

We identify interactions with stakeholders in different stages of risk assessment and communication in which stakeholders hold passive or active positions in the contribution of information. Interaction with risk assessment organizations can appear in two different forms: the inclusion of data or information on stakeholders but also the provision of (not necessarily stakeholder-related) evidence or criticism by stakeholders. Those interactions are partly defined in detail, partly more by their implicit nature. For example, the active influence of consumers in the process of risk assessment is partly indirect and difficult to measure. We find controversial opinions by experts on its scope.

3.3.6.1. Differences between cases

Nevertheless, as expected in a maximum variation sampling strategy, the case studies differed partly on the sub-theme level[39]. Possibilities for stakeholder interaction differ due to the nature of hazards and thus the structure of the risk assessment (IoM & NRC,

[39] Sub-themes which only occurred in one case are marked in Appendix D.

2003). We find more food sector-related variables in microbiological risk assessment, especially in terms of food production. Additionally, consumer habits were only relevant for microbiological risks. This can be explained by the relevance of food production and preparation for the dynamic growth of microbes. For plant protection products, experts reported regulatory frames for used data, quality criteria or variables. The intentional introduction of these substances allows for a definition of those in advance. The differences between cases support the assumption that the trust, role and possibilities for interaction of stakeholders in risk assessment organizations are case specific. Although we can provide a cross-case description, it is important to note that processes and external conditions differ due to the nature of hazard. Therefore, detailed processes within the theoretical description need to be described case by case.

3.3.6.2. Types of information on stakeholders

For food sector stakeholders, we find a variety of information used in risk assessment organizations (more variables in microbiological risk assessments). For consumer stakeholders, results confirm that we find different types of information from social sciences like acceptance, trust, or perception. This information is mostly used in risk communication but has indirect effects on risk assessment. For example, risk communication identifies emerging risks from risk screenings that are then scrutinized in risk assessment. Our results expand on previous descriptions that ignore the use of this information exchange in risk assessments. The non-explicit nature of these interactions supports the need for a stronger emphasis of the exchange of societal evidence between assessment and communication. Risk communication plays a central role in producing and collecting societal evidence on risks. It remains an interesting question if it would be possible to integrate this information more systematically, i.e., in structured approaches to risk assessment organizations. This might emphasize already-existing efforts to include stakeholder evidence and to defend the social robustness of assessments.

We do not find a use of information on benefits, risk equality, economic or welfare losses, neither for food sector nor for consumers. This type of evidence is commonly used in risk management and it might be asked, who is responsible for providing this type of non-health-related information and who should conduct risk evaluations (Dreyer & Renn, 2009; König et al., 2010; Ruzante et al., 2010). It is necessary to discuss whether it might

be possible that risk assessment provides all necessary and relevant information for decisions (health, economic considerations, societal impacts). This is especially important if we see risk assessment as a "rational and objective picture [of] what is known or believed to be known" (FAO & WHO, 2009) and the source of "all available information" (FAO & WHO, 1997). Solid evidence from social sciences can contribute to the conceptualization of acceptable risk levels (Fischhoff, 1994). Structured scientific approaches, specified sets of quality criteria and clear responsibility allocations for these types of information might increase transparency and stability of risk assessments and the resulting regulations (FAO & WHO, 2009; van Zwanenberg & Millstone, 2000; Wentholt et al., 2009).

3.3.6.3. Interaction mechanisms and (potential) roles of stakeholders

In selecting and using evidence, we identify different types of stakeholder interaction. Hereby, stakeholders have different roles: They act as providers of information, study objects, discussion partners, influencers, criticizers, test objects or information recipients.

On the interaction level, it is necessary to acknowledge the variety of possible interactions. Different phases in risk assessment processes entail different possibilities for interaction. Our study shows that it is important to specifically define which participation measures are targeted to allow effective research, evaluation, and participation processes. It extends the previous literature on participation. In the cases covered, this includes many forms of activities: Forms of communication (one-way), consultation (one-way), or participation (two-way) (Rowe & Frewer, 2005). While Rowe & Frewer (2005) define engagement forms based on information flow, we define them based on the direction of force for the initiative, which also includes information flows. This adaptation enabled us to give initiating agency to stakeholders and identify additional interactions which extend these classification scheme: social media allows people to react on communication (two-way) or communication tools are evaluated by the public (two-way). In public consultation, we do not only find a formal dialogue but more the collection of (representative) information. In participation, we find an additional category: Active forms of participation initiated by stakeholders. Additionally, it is important to note that there are indirect forms of interactions, for example, the exchange of information on stakeholders within the risk analysis process or contextual public debates on certain

topics. According to our results, it would be possible to label many more interaction forms as participatory actions. This could contribute to defending risk assessments against criticism of being an elite action. Clearly identifying interactions increases transparency, allows for specific definition and control of efficiency in all processes (Rowe & Frewer, 2005).

This leads us to the question which roles stakeholders can assume during the process of risk assessment. The experts seem to support general objectives of public participation in risk assessment, namely, to satisfy regulatory requirements, resolve conflicting views, increase transparency/defensibility, (ex)change views, improve services, determine needs, empower citizens, support/foster trust and credibility or elicit values (Homan et al., 2001). Experts expressed, for example, concerns regarding trust and mutual understanding. Nevertheless, it is important to not reduce participation to a set of outreach activities (Devos et al., 2019). Literature offers additional objectives of participation, for example, participation in framing, comprehensive scoping of problems, including local information, mutually discussing alternatives, using the public as quality assurance, identifying relevant sub-groups, or determining appropriate standards and criteria. For these objectives, mechanisms are not well defined in most cases (Barker et al., 2010; Homan et al., 2001; Klinke & Renn, 2002; Millstone, 2009). We find some of these elements in our sub-themes in the form of defined processes (risk screenings to identify new risk topics) or ad-hoc phenomena (role of stakeholders as critics and quality assurance). Experts partly value outside criticism as a control mechanism, but also reported indicated disappointment in mistrust. Framing is another possibility for stakeholder inclusion because quantitative risk assessments do not “fully encompass the complexities that are included in many stakeholder beliefs” (Barker et al., 2010). The systematic identification of contextual and normative ambiguities can support the integration of different perspectives of different stakeholders (Johansen & Rausand, 2015; Klinke & Renn, 2002). In academia, new approaches have been developed to bring more social and economic information into risk analysis, for example, through complete reorganization of regulatory frameworks (Dreyer & Renn, 2009; König et al., 2010), specific risk prioritization schemes (Fischhoff & Hope, 1984; Ruzante et al., 2010), or new models of stakeholder involvement (Devos et al., 2019). These approaches acknowledge that food safety is a field that depends on contextual information.

Stakeholders such as food producers are needed to describe the food safety problem and to provide information on practice.

Generally, participation makes sense in certain stages but may not be useful in others. Engagement needs to match context (Rowe & Frewer, 2005). Otherwise, it causes frustration for experts: Participation mechanisms cause additional workload for risk assessment organizations, which can be seen as problematic or impractical (Barker et al., 2010; Devos et al., 2019). This view was supported by experts who reported missing resources through and for answering consumer concerns. Participation might delay processes, and is problematic especially in time-critical cases (Barker et al., 2010). Problems of participatory information are extensively discussed in the literature, for example, that beliefs can change rapidly or that participant experience can vary (Barker et al., 2010; Gruev-Vintila & Rouquette, 2007).

Besides participation, independence of risk assessment must be maintained (T. Robinson et al., 2016). An independent risk assessment is a key element of European food safety regulations (EC & EP, 2002). Experts reported this as critical element in stakeholder participation, especially for the food sector with economic interests. Independence was also formulated as a quality criterion. This might indicate a fundamental conflict between stakeholder participation and independence. A possible solution is the invention and evaluation of consistent and transparent processes for stakeholder participation, defined for each stage of the information flow also for information which does not fit into current frameworks. An important note is that "participatory information does not replace or modify information from other sources but becomes part of an extended domain" (Barker et al., 2010). Future research might concentrate on the specific links between interactions and their influence in risk assessments.

The findings suggest that it is important to understand that different stakeholders have different resources and face different barriers. Therefore, it might be that, for example, relevant evidence from consumers is not included because they have no possibility of bringing relevant information to the process (Fischhoff, 1994; Hassauer & Roosen, 2020; Hirschman, 1971; Kinchy, 2010; Sunstein, 2018). For example, experts reported that they can only include scientific information in the input section. This means that possibilities for participation vary between expert stakeholders and lay stakeholders (Barker et al., 2010). Especially information from lay people can have different sources and different

quality (Fischhoff & Hope, 1984). Therefore, quality criteria act as filters or barriers that have to be passed by stakeholders to provide input (Vogel & Delfini, 2008). This might be hard to achieve for consumers because they lack technical knowledge or financial resources (Hassauer & Roosen, 2020; Kinchy, 2010; Sunstein, 2018). In general, it is necessary to strengthen interactions between risk assessment organizations and stakeholders to increase the relevance of risk assessments. It will also be necessary to define an eco-system of sharing of information produced by the different stakeholders, respecting the confidentiality, and data on production and management of food safety including data on monitoring and control and traceability of food. Additionally, the representation of context-related quality criteria confirms that quality in risk assessments is contextual and not necessarily absolute (Funtowicz & Ravetz, 1993). On the other hand, quality criteria are needed as a stabilizing mechanism against destabilizations from context or the stakeholders themselves (van Zwanenberg & Millstone, 2000). It is important to discuss whether it is necessary to define quality criteria for in- and exclusion of non-scientific information. Again, risk assessment organizations face the conflict regarding how to lower participation barriers and to keep independence and objectivity.

3.3.6.4. Risk assessors between societal and scientific expectations

Consequently, the question of participation affects members of risk assessment organizations and especially the group of risk assessors in different ways. First, they stand between regulations, societal expectations, and a representation of scientific values (Johansen & Rausand, 2015; Vareman & Persson, 2010). It is difficult for risk assessors to deal with these conflicts and the consequential normative ambiguity (Johansen & Rausand, 2015; Stirling, 2008). “It is typically unrealistic to think that these [societal] influences can be eliminated or ignored. This often prevents the necessary reflection about values and can result in conflict and distrust” (Devos et al., 2019). Our results confirm these tensions and show that risk assessors are directly affected by them. Second, risk assessors face external limits through regulatory framing. For example, the scopes of chemical approvals are determined in regulations (Vareman & Persson, 2010). Experts report for example in context-related quality criteria that this type of tight regulatory framework frames the scope of their work. Additionally, they report for example that opportunities to collect specific data depend on financial resources. Third, risk assessors

face the issue that they are not responsible for framing risk assessments. They provide answers to questions posed by risk management, which is problematic if information from and on stakeholders is relevant but does not fit into these framings in terms of quality and relevance. It remains an open question as to who is responsible for including these types of evidence—risk assessment or risk management. This is a fundamental problem because if risk management frames risk assessments, it influences science per se, introduces external limits to sciences, and thus threatens independence. Contrary, if risk assessment would frame itself, it would have to make value decisions, which is not its task (Vareman & Persson, 2010). The same difficulty arises if political framing is insufficient. We find these conflicts in the statements of experts, for example, in the context-related quality criteria. Experts state that they are criticized for processes that they are not responsible for. It might be helpful to define processes for stakeholder interaction that could be used in framing. Risk assessment might provide the information, risk management the decisions on framing.

3.3.6.5. Limitations

The study has various limitations that need to be considered when interpreting and using its results. Limitations related to the data collection include the small sample size. This can be explained by the small population of experts with the required characteristics. Unique insights in risk assessment processes, the sampling strategy, and high information density in the interviews support the assumption that we got close to a level of theoretical saturation (Patton, 2015). The choice of experts by the organizations was necessary because publicly available information is limited; however, this might introduce a bias to our study. Due to circumstances, we conducted a part of the interviews not in person but on the phone and cannot exclude an impact of the interview settings for interview flow, interactions, or connection to the interviewer. It might be problematic that interview guidelines were send in advance. This was necessary because the interview questions were complex and required a certain amount of preparation. We try to reduce this bias by sending only the part of the interview guide in advance which requires preparation.

Another potential bias might arise from the definition of experts. In general, experts are not seen as providers of objective knowledge but as providers of experiences and process knowledge. The theoretical description of cases is based on the selected experts (Meuser

& Nagel, 2009). Therefore, we are aware that sub-themes (see Appendix D) might be incomplete. Sub-themes rather serve as examples to illustrate the theoretical description of the cases.

As described in the introduction, stakeholder groups consist of different actors, the food sector of primary production, food processing etc. and the consumer group of individual consumers and consumer associations. We are aware that different actors might participate differently in interaction mechanisms. Nevertheless, we are not able to discuss these roles in this paper explicitly because it aims to identify basic interaction types across the stakeholder groups. Therefore, we are not able to describe all interactions in the whole process.

Our results identify two main areas of interest for the discussion: Used types of information on stakeholders and interactions of risk assessment organizations with stakeholders in the determination of evidence. To discuss the results and implications, we use specific examples which are partly not mentioned in the result section due to complexity. All examples are derived from sub-themes listed in Appendix D.

3.3.7. Conclusions

Results confirm interactions of risk assessment organizations with stakeholder groups in different directions (active participation, passive integration, or exchange) on different levels (input, processing, output). The theoretical description identifies complex interactions with stakeholders in different stages of research processes (van Zwanenberg & Millstone, 2000). Our results highlight that the determination of evidence for food safety cannot be seen as independent of stakeholders. Stakeholders define risks themselves and influence the relevance of these risks. This contradicts a purely realist perspective on risk assessment, which would negate these complex interactions (Devos et al., 2019). Results support the allegation that there is a need to recognize these influences and a way to make them transparent for stakeholders as well as members of risk assessment organizations because they are partly rather implicit. This means that processes must be defined for active and passive integration of stakeholder information and are not limited to framing. Including more real-world complexity in the form of stakeholder evidence might increase the stability of risk assessments as social

organizations (van Zwanenberg & Millstone, 2000). Additionally, a reasonable inclusion of stakeholders acknowledges and considers the societal ambiguity of risk issues that also affect risk assessment organizations (Johansen & Rausand, 2015). We emphasize the need for transparent and structured analysis and definition of interaction processes and targeted outcomes in every stage of information flow and for separate stakeholders.

3.3.8. Acknowledgements

This article is based on research supported by funding from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Research Group 2448.

3.4. Bewerten und Gewichten: Evidenz als Entscheidungshilfe in der Gesundheits- und Umweltpolitik (English translation: Evaluating and weighting: Evidence as decicion tool in health and environmental policy)

Publication: Haßauer C., Ehlers S., Roosen J. (2019). Bewerten und Gewichten: Evidenz als Entscheidungshilfe in der Gesundheits- und Umweltpolitik. In Zachmann, K., & Ehlers, S. (2019). Wissen und Begründen – Evidenz als umkämpfte Ressource in der Wissensgesellschaft.

3.4.1. Abstract

Evidence is a central element in regulations in plural societies. In regulations, it is necessary to negotiate various interests and effects on different groups. Complex decision settings require a justification of governmental interventions. Thus, different procedures, so called evidence practices, have been established to anticipate effects of decisions using standardized indicators. One example for such evidence practices is the cost-benefit analysis as example for an established economic procedure. It is used to evaluate and weight different forms of evidence based on a monetarization of effects. The objective of this essay is to examine the evidence practices of evaluation and weighting using the example of cost-benefit analysis. The essay provides an analysis of cost-benefit analysis, including its mechanisms, its establishment in policy, and a discussion of its critical issues. That contributes to a reflective and critical view on the evidence practices of evaluating and weighting evidence as decision tool in policy. This might support a contextual understanding and interpretation of the results of such analysis. The essay shows the evidence practices of interest as instruments of power in policy and tools in reasoning and conflict management.

3.4.2. Einleitung

Sind wir bereit aus Profitgründen […] Artensterben, Bodendegradierung und nitrithaltiges Trinkwasser zu akzeptieren? (Initiative Volksbegehren Artenvielfalt, 2019)

Evidenz als Entscheidungshilfe ist ein zentrales Element für Regulierungen und Interventionen in pluralen Gesellschaften. Öffentliche Entscheidungen sind komplexe

Vorgänge, schließlich gilt es nicht nur eine Vielzahl möglicher Folgen, sondern auch unterschiedliche Interessen zu berücksichtigen. Insbesondere staatliche Eingriffe, die die Freiheit von Einzelnen einschränken, sind in demokratischen Gesellschaften zu begründen. Dafür haben sich Verfahren – von uns im Folgenden als Evidenzpraktiken verstanden – etabliert, mit denen die Auswirkungen von Entscheidungen in systematisierter Form mithilfe von Messgrößen und formalisierter Kriterien antizipiert werden[40]. Insbesondere die Kosten-Nutzen-Analyse konnte sich als zentrale Praxis ökonomisierter Entscheidungsprozesse durchsetzen, anhand derer die Folgen staatlicher Eingriffe bewertet und gewichtet werden. Dieses Kapitel wird diese Bewertungs- und Gewichtungsprozesse als Element der evidenzbasierten Politik betrachten. Dazu werden wir immer wieder auf zwei Fallstudien rekurrieren, die Regulierung von Pestiziden sowie international gesteuerte Ansätze zur globalen Bekämpfung von Malaria, um die Praktiken zu veranschaulichen.

Gerade in einer Zeit des wachsenden Einflusses sogenannter „postfaktischer" Geltungsbehauptungen, des Erstarkens der Populisten und der zugehörigen Kommunikationsstrategien in den Medien verspricht der Bezug auf Evidenz, also die Mobilisierung von gesellschaftlich anerkannten Daten und Fakten, auf eine gemeinsame Entscheidungsbasis zurückzukehren. Durch Evidenz lassen sich folglich Entscheidungen zu kontroversen Themen rechtfertigen und Entscheidungsprozesse rationalisieren (Dobrow et al., 2004). Von evidenzbasierter Politik wird erwartet, dass sie Ideologien und Machtasymmetrien in der Entscheidungsfindung neutralisiert (Lancaster et al., 2017). Gleichzeitig verspricht die wissensbasierte Legitimation politischer Maßnahmen eine Erhöhung der Wirksamkeit dieser Maßnahmen sowie eine bessere Ressourcennutzung und eine geringere Fehleranfälligkeit (Evidence-Based Policymaking Collaborative, 2019).

Ein zentrales und kritisches Element der evidenzbasierten Entscheidung ist die Abwägung zwischen verschiedenen Auswirkungen und Interessen. So betreffen

[40] In der Europäischen Union sind zum Beispiel Standardindikatoren für evidenzbasierte Entscheidungen in den Bereichen nachhaltiges Wachstum, Sozialpolitik oder Währung definiert. Auch die Millenium Development Goals der Vereinten Nationen basieren auf einer Sammlung statistischer Indikatoren, anhand derer die evidenzbasierte Umsetzung von armutsbekämpfenden und auf die Nachhaltigkeit bezogenen Maßnahmen analysiert werden können (Statistisches Bundesamt, 2019).

politische Entscheidungen nicht nur einen Aspekt in der Lebensumwelt der Menschen, sondern ganz unterschiedliche Bereiche. Die Zulassung eines Pestizids wirkt sich nicht nur auf die Gesundheit von Anwendern und Konsumenten aus, sondern auch auf die Biodiversität, Wirtschaft und Verteilung von Vermögen und Umweltresourcen. Dabei sind diese Auswirkungen für verschiedene Akteure ungleich relevant, wie sich auch in Debatten um Pestizidzulassungen zeigt: Wenn beispielsweise der Einsatz eines Pestizids mit dem Verschwinden bestimmter Arten, aber auch mit der Steigerung der landwirtschaftlichen Produktivität einhergeht, ist davon auszugehen, dass seine Zulassung von Naturschutzgruppen und Landwirten unterschiedlich beurteilt wird. Um evidenzbasiert demokratische Entscheidungen treffen zu können, sind solche Auswirkungen daher nicht nur zu messen oder zu ermessen[41], sondern sie sind auch zu *bewerten* und zu *gewichten*, insbesondere da an diese Entscheidungen oft auch die Verteilung von Ressourcen geknüpft ist. Damit werden *Bewertung* und *Gewichtung* zur Basis für den Austausch von Argumenten und zu wichtigen Bestandteilen im Modus der Konfliktbewältigung.

Im Folgenden werden die Evidenzpraktiken *Bewerten* und *Gewichten*, die der Findung und Begründung von Entscheidungen dienen, kritisch diskutiert. Dabei wendet sich das Kapitel im zweiten Abschnitt der Beschreibung der ökonomischen Praktik der Monetarisierung von Effekten (*Bewertung*) sowie deren Abwägung (*Gewichtung*) im Rahmen der Kosten-Nutzen-Analyse zu. Dabei wird ausgeführt, wie in der Kosten-Nutzen-Analyse zum Beispiel gesundheitliche Auswirkungen in monetäre Werte überführt und gegen wirtschaftliche Ertragsverluste gewichtet werden können, wodurch eine neue Form von (ökonomischer) Evidenz entsteht.

Im dritten Abschnitt betrachtet das Kapitel die Etablierung der Kosten-Nutzen-Analyse als formalisierte Form der Bewertung und Gewichtung in Entscheidungsprozessen. Die Kosten-Nutzen-Analyse trug maßgeblich dazu bei, dass sich die Praktiken von Bewertung und Gewichtung langsam vom nicht kodifizierten „Ausdiskutieren“ mit unklaren Verantwortlichkeiten zu einem idealerweise transparenten, nachvollziehbaren Evaluierungs- und Entscheidungsprozess entwickelte. Diese Entwicklung geschah vor

[41] Eine ausführlichere Abhandlung dieser Praktiken ist im Kapitel *Messen und Ermessen* dieses Buches zu finden (Zachmann & Ehlers, 2019).

dem Hintergrund gewachsener gesellschaftlicher Ansprüche an politische Transparenz sowie der Globalisierung, die die Harmonisierung und Vergleichbarkeit politischer Maßnahmen erfordert.

Der vierte Abschnitt zeigt beispielhaft an der *Roll-Back-Malaria*-Initiative, wie ökonomisierte Entscheidungsprozesse internationaler Organisationen und insbesondere die Kosten-Nutzen-Analyse in einem Politikfeld wirken, das entscheidend von konkreten, lokalen Bedingungen globaler Armut geprägt ist. Im Mittelpunkt steht dabei die lokale Anwendung internationaler Vorgaben und den Problemen kontextspezifischer Evidenz. Auch wird die Übertragung allgemeiner Parameter und Determinanten auf die verarmten Regionen der Welt diskutiert.

Zuletzt sollen kritische Aspekte der Kosten-Nutzen-Analyse diskutiert werden. Dies ermöglicht eine differenziertere Betrachtung der Evidenzpraktiken *Bewerten* und *Gewichten.* Weiterhin werden die Evidenzpraktiken in einen größeren Zusammenhang gestellt und es wird die Frage aufgeworfen, inwieweit durch diese Verfahren Transparenz geschaffen wird – und nicht zuletzt, inwieweit damit Evidenz, verstanden als eine Qualität von Wissen, in Entscheidungsprozessen erzeugt und nutzbar gemacht werden kann.

3.4.3. Bewerten und Gewichten als Evidenzpraktiken zur Begründung von Entscheidungen

Bewertungen und *Gewichtungen* sind vor allem dann relevant (aber auch umstritten), wenn Entscheidungen komplexe Effekte auf unterschiedliche Akteure implizieren. Am im Folgenden diskutierten Fallbeispiel, der Zulassung von Pestiziden, lassen sich gerade aufgrund der Komplexität von Effekten die Praktiken *Bewerten* und *Gewichten* besonders deutlich illustrieren: Entscheidungen in der Pestizidregulierung können eine Vielzahl von ökologischen, ökonomischen und sozialen Konsequenzen aufweisen, die in ihren einzelnen Dimensionen zu *bewerten* und zu *gewichten* sind. Ein großes Spektrum von Pestiziden dient in der Produktion von Nahrungsmitteln einerseits der Erhöhung der Produktivität und Qualität, sorgt für die Vermeidung von Resistenzen und vermindert das Produktionsrisiko. Andererseits verschmutzen Pestizide die Umwelt und bergen ein Gesundheitsrisiko für Anwender und Konsumenten (Sexton et al., 2007). Beide Evidenzpraktiken, *Bewerten* und *Gewichten,* dienen der Darstellung und im Zweifel der Rechtfertigung, warum bestimmte Entscheidungen, zum Beispiel die Zulassung von

Pestiziden, getroffen werden und beschreiben daher Modi des Begründens mit dem Ziel der Konfliktbewältigung. Die Praktiken sind hierbei nicht als Gegensatzpaar zu sehen, sondern ergänzen sich: In der Zulassung eines Pestizids muss einer Gewichtung der Evidenz zu den Wirkungen in unterschiedlichen Dimensionen, wie zum Beispiel der ökologischen oder gesundheitlichen Folgen, immer eine Bewertung der einzelnen Folgen vorangehen. Eine umfassende Beurteilung eines Pestizids ist meist nur durch eine Gewichtung verschiedener Konsequenzen möglich.

Die Praktik des *Bewertens* nimmt eine Wertzuschreibung für einen einzelnen Effekt einer Entscheidung vor. Dies geschieht in Kosten-Nutzen-Analysen durch Monetarisierung, also die Zuschreibung eines Geldwertes. Dazu stehen unterschiedliche Methoden der ökonomischen Bewertung zur Verfügung. Während für Marktgüter wie Getreide Marktpreise für die Zuschreibung eines Wertes herangezogen werden können, verlangen Umweltressourcen wie Böden oder Gesundheit nach anderen Bewertungsverfahren. Beispielsweise wird in der Bewertung von Leben und Gesundheit der Wert eines statistischen Lebens (*Value per Statistical Life*, VSL) als Bewertungseinheit herangezogen. Somit geht es also nicht um die Gefährdung eines identifizierten individuellen Lebens, sondern um die Veränderung des Krankheitsrisikos oder der Sterberate. Die damit einhergehenden volkswirtschaftlichen Kosten können einerseits durch den Humankapitalansatz, der gesundheitliche Folgen auf Basis des Wertes der Produktivität eines Individuums berechnet, oder die Messung der Zahlungsbereitschaft für Nichtmarktgüter [42] beziffert werden. Dass diese Berechnungen in der Praxis keineswegs unumstritten sind, zeigt die später folgende Analyse der Bewertungsverfahren von statistischen Leben in der *Roll-Back-Malaria*-Initiative.

Bewertungen können also monetäre Wertzuschreibungen wie Getreidepreise, Produktivitätsausfall durch gesundheitliche Folgen des Pestizideinsatzes für die Anwender oder aber auch die Zahlungsbereitschaft von Bürgern für die Vermeidung des Aussterbens einer Insektenart sein. Dabei zeigen Studien, dass auch monetäre Bewertungen Über- oder Unterschätzungen aufweisen können. So unterschieden sich bei der Schätzung des Wertes eines statistischen Lebens im Zusammenhang mit

[42] Zahlungsbereitschaften werden auf Basis der Messung von Präferenzen erhoben (Maria Travisi et al., 2006).

Pestizidregulierungen die Höhe des angesetzten Wertes zwischen Anwender- (35 Millionen US-Dollar) und Konsumentenleben (60.000 US-Dollar) (Cropper et al., 1992). Auch die Zahlungsbereitschaft von Konsumenten für die Reduzierung von Pestizidrisiken variiert stark abhängig vom Untersuchungsdesign und dem Ein- oder Ausschließen von bestimmten Risiken (Florax et al., 2005).

Bewerten verweist weiterhin auf eine qualitative und sogar normative Dimension: Zunächst ist es für eine Regulierung notwendig zu bestimmen, welche Effekte bei der Entscheidungsfindung berücksichtigt werden. Pestizide können bestimmte Langzeitfolgen für Ökosysteme, die menschliche Gesundheit und landwirtschaftliche Praxis haben. Eine normative Form der *Bewertung* ist demnach, ob diese Faktoren in der Entscheidungsfindung berücksichtigt oder ignoriert werden. Dies scheint für manche Effekte, wie beispielsweise eine nachgewiesene Häufung von Krebserkrankungen, selbstverständlich, kann aber insbesondere bei unsicheren, indirekten oder weit in der Zukunft liegenden Folgen für Umwelt oder Gesundheit eine kritische und keinesfalls einfache Entscheidung sein.

Eine weitere Form der *Bewertung* kann die Einstufung eines Faktors als Ausschlusskriterium sein: Die Klassifizierung eines Pestizids als mutagen bedeutet in der EU automatisch ein Verbot der Substanz, ohne Berücksichtigung weiterer Kosten oder Nutzen. Ein *Gewichten* unterschiedlicher Konsequenzen wird also explizit ausgeschlossen.

Die Praktik des *Gewichtens* bezieht die Bewertung der einzelnen Effekte mit ein und rückt die Abwägung zwischen einzelnen Konsequenzen von Entscheidungen in einem komplexen System durch eine Quantifizierung der zugeordneten Relevanzen stärker in den Blick (Merriam-Webster, 2019). Dies ist insbesondere deshalb wichtig, da Trade-Offs in komplexen Entscheidungsfindungen wahrscheinlich sind. In diesen Entscheidungen ist die Erreichung verschiedener Ziele unvereinbar, das heißt die Verbesserung eines Ziels ist nur mit einer einhergehenden Einschränkung anderer Ziele möglich. Dabei können Trade-Offs unterschiedlich klassifiziert werden, je nachdem ob gegenläufige Risiken (selbe oder unterschiedliche) entstehen oder andere Bevölkerungsgruppen betroffenen sind (eine oder unterschiedliche) (Gray & Hammitt, 2000). So müssen zum Beispiel bei der Zulassung von Pestiziden gesundheitliche, ökonomische und umweltbezogene Auswirkungen einbezogen und gegeneinander

abgewogen werden. Die Zulassung eines neuen Pestizids kann Resistenzbildungen von Schadorganismen vorbeugen oder sicherer in der Anwendung sein, auf der anderen Seite aber bestimmte Nichtzielorganismen schädigen (Sexton et al., 2007). Dabei stellt sich die Frage, welches Ziel wichtiger ist: Die Vermeidung der Resistenzbildung oder der Schutz von Nichtzielorganismen?

Insbesondere die Bestimmung von gesellschaftlich akzeptablen Wirkungen oder Grenzwerten erfordert eine Auseinandersetzung mit den Effekten der Maßnahmen auf die konkret davon betroffenen Akteure (Fraiberg & Trebilcock, 1998). Diese Aufgabe ist herausfordernd, da Effekte transnational und generationsübergreifend sein können und häufig moralische Fragestellungen aufwerfen. Pestizide können sich in der Nahrungskette anreichern und über Ländergrenzen hinweg verbreiten. Ihre Speicherung in Böden kann Jahrzehnte überdauern und dadurch auch das Wohlergehen zukünftiger Generationen beeinflussen. Die Bewertung der weitreichenden Effekte kann damit über die Wirkung auf die gegenwärtige Gesellschaft hinausgehen (Taebi, 2017). Damit sind in diesen Fällen weder eine klassische, wissenschaftliche Risikobewertung noch die Untersuchung gesellschaftlicher Akzeptanz als alleinige Instrumente geeignet, um komplexe Entscheidungen zu rechtfertigen (Wagner, 2016). Die Kosten-Nutzen-Analyse bietet ein Instrument, um diese Lücke zu schließen und unterschiedlichste Folgen auf einer Argumentationsebene zu *bewerten* und zu *gewichten*. Anhand der Kosten-Nutzen-Analyse in der Pestizidzulassung soll im Folgenden aufgezeigt werden, wie sich die Evidenzpraktiken der *Bewertungen* und *Gewichtungen* aus der ökonomischen Theorie entwickelten, zunehmend formalisiert, implementiert und so zu einem festen Bestandteil evidenzbasierter Politik wurden. Dies erleichtert eine kritische Einordnung der Evidenz, die in Kosten-Nutzen-Analysen erzeugt wird.

3.4.4. Die Etablierung der Kosten-Nutzen-Analyse in komplexen Entscheidungsprozessen

Bei der *Bewertung* evidenzierter Auswirkungen spielen ökonomische Konzepte eine große Rolle. Seit der Entstehung in den 1930er Jahren in den USA entwickelte sich die Kosten-Nutzen-Analyse als eine etablierte Form der formalisierten *Bewertung* und *Gewichtung* der Auswirkungen von Regulierungen und Maßnahmen, die seither als

Evidenz für Entscheidungsbegründungen eingesetzt wird[43]. In der Kosten-Nutzen-Analyse gelten gute Entscheidungen als solche, bei denen der Nutzen für die Gesellschaft insgesamt größer ist, als die Kosten es sind. Folglich werden etwa gesundheitliche, ökologische und ökonomische Effekte der Zulassung eines Pestizids in Form von Kosten und Nutzen direkt gegenübergestellt, die vorher monetär *bewertet* wurden.

Für die großen Vorhaben in der New-Deal-Politik unter Roosevelt rechtfertigte der US Corps of Army die Projekte durch eine auf einem Investitionskalkül basierende Kosten-Nutzen-Berechnung. Dass diesen Berechnungen wohlfahrtsökonomische Bedeutung zugeschrieben werden konnte, zeigten die Aufsätze von Kaldor und Hicks 1939 (Hicks, 1939; Kaldor, 1939). Das von ihnen eingeführte Kaldor-Hicks-Kriterium einer potentiellen Paretoverbesserung[44] wurde in der Folge zentraler Bestandteil von Kosten-Nutzen-Analysen. Es besagt, dass Veränderungen in der wirtschaftlichen Ordnung nicht notwendigerweise tatsächlich alle betroffenen Parteien besserstellen müssen, sondern dass es ausreicht, wenn eine solche Besserstellung von betroffenen Personen durch Umverteilung potentiell möglich ist. Durch Kosten-Nutzen-Analysen, die das Kaldor-Hicks-Kriterium anwenden, können Eingriffe also auch dann gerechtfertigt werden, wenn die Gewinner die Verlierer für ihre Verluste potentiell kompensieren können. In diesem Fall können Verluste – zumindest theoretisch – ausgeglichen werden. Das Prinzip der potentiellen Paretoverbesserung erlaubt es also, die Praktiken des *Bewertens* und *Gewichtens* zur Gewinnung von Handlungsempfehlungen zu verwenden, selbst wenn die entsprechenden Eingriffe auch negative Konsequenzen haben. Allerdings müssen die „Gewinne“ die Verluste nur potentiell kompensieren, was mit einer Ignoranz gegenüber möglichen realen Verteilungswirkungen einhergehen kann. Überträgt man dieses theoretische Argument auf das Verbot von Pestiziden, kann also eine Reduzierung des Gesundheitsrisikos für Verbraucher (Nutzen) eine Steigerung des Ertragsverlustrisikos (Kosten) rechtfertigen: Obwohl das Verbot eine Gruppe schlechterstellt (Landwirte), bringt diese Entscheidung insgesamt einen Netto-Nutzen (Cropper et al., 1992; Gray & Hammitt, 2000).

[43] Zur Geschichte der Kosten-Nutzen-Analyse vgl. Porter (1997).

[44] Nach dem Pareto-Kriterium ist eine Maßnahme dann zu befürworten, wenn die Wohlfahrt mancher Gesellschaftsmitglieder verbessert wird, ohne dass jemand schlechtergestellt wird.

Theoretisch basiert die Kosten-Nutzen-Analyse auf der neoklassizistischen Annahme eines vollkommenen Markts, eine Annahme, die selbst in den Wirtschaftswissenschaften höchst umstritten ist. Auch macht eine Kosten-Nutzen-Analyse eine Monetarisierung von allen Kosten und Nutzen notwendig (Driesen, 2004), welche eine Reihe methodischer und moralischer Fragen mit sich bringt, vor allem in der Bewertung von Effekten etwa für Umwelt oder Menschenleben, denen im Markt kein Preis zugeschrieben wird (Fraiberg & Trebilcock, 1998). Es ist also festzuhalten, dass die mit Kosten-Nutzen-Analysen erzeugte Evidenz überaus fragil ist[45]. So mag auch zu erklären sein, warum sich die Anwendung von Kosten-Nutzen-Analysen in der Umweltpolitik nicht als alleiniges Instrument durchsetzen konnte: In den USA sind Kosten-Nutzen-Analysen zum Beispiel im „Clean Water Act" verboten, in der Regulierung von verschiedenen Pestiziden aber vorgeschrieben (Cropper et al., 1992). Alternative Ansätze in der Umwelt- und Gesundheitspolitik bieten zum Beispiel Risikoanalysen oder die Bewertung von Risk-Trade-Offs, in denen der Schritt der Monetarisierung nicht notwendig ist.[46] Akzeptable Grenzwerte werden in diesen Verfahren nicht durch Netto-Nutzen berechnet, sondern durch naturwissenschaftlich determinierte Grenzwerte oder Bewertungen mit anderen Indikatoren, wie den qualitätskorrigierten Lebensjahren für Gesundheitseffekte (Gray & Hammitt, 2000). In diesen Verfahren fehlt aber die Möglichkeit zum direkten Vergleich beispielsweise von Umwelteffekten und Anwenderschutz.

Bewerten und *Gewichten* und ihre Umsetzung im Rahmen der Kosten-Nutzen-Analyse bedingen somit eine Vielzahl normativer Entscheidungen, um eine Wertzuschreibungen zu ermöglichen. Diese Wertzuschreibung hängt immer vom jeweiligen Kontext, z.B. von der Verteilung von Nutzen, Kosten und Risiken in der Gesellschaft, ab. Idealerweise erfolgt die Wertzuschreibung im Kontext der Gesellschaft, für die Entscheidungen zu treffen sind. Nur dann kann davon ausgegangen werden, dass die aus einer Kosten-Nutzen-Analyse resultierenden Empfehlungen die Werturteile eben dieser Gesellschaft widerspiegeln.

[45] Mit dieser Kritik setzen wir uns im späteren Verlauf des Kapitels noch ausführlicher auseinander.

[46] Genauere Erläuterungen zu naturwissenschaftlichen Verfahren in der Risikoanalyse sind im Kapitel *Messen und Ermessen* dieses Bandes zu finden.

Ökonomen betrachten die Kosten-Nutzen-Analyse trotz ihrer Limitierungen dennoch als hilfreiches Instrument zur Rationalisierung von Regulierungsentscheidungen, die auf die Beförderung der gemeinsamen Wohlfahrt (Arrow et al., 1996) ausgerichtet sein sollen. Die Kosten-Nutzen-Analyse bietet generell eine Möglichkeit zur Systematisierung und Disziplinierung von komplexen politischen Entscheidungen (Fraiberg & Trebilcock, 1998). Sie schafft weiterhin eine Basis und gemeinsame Währung für den Austausch von Argumenten und die Abschäzung und Beurteilung zukünftig auftretender Auswirkungen (Lynch, 2012). Kosten-Nutzen-Kalküle werden inzwischen in unterschiedlichsten Kontexten und Staaten diskutiert und angewendet, so seit Anfang der 2000er Jahre auch in der Politik der EU. Beschleunigt wurde die Implementierung unter anderem in den 1980er und 1990er Jahren, in denen nach verschiedenen Krisen die Rolle von Wissenschaftlern und die Transparenz in Entscheidungen hinterfragt und somit die Trennung von Wissenschaft und Entscheidung forciert wurde [47]. Dies hing auch zusammen mit den zunehmenden Diskursen um Risiko, Unsicherheit und Nichtwissen. Weiterhin spielte bei der Verbreitung der Kosten-Nutzen-Analyse auch die Globalisierung eine Rolle, da ein freier Handel und universal anwendbare (Entwicklungs-) programme eine vergleichbare Regulierungspraxis und damit auch Bewertungsmethoden erforderlich macht.

3.4.5. Evidenz in der Roll-Back-Malaria-Initiative

Globale Gesundheits- und Entwicklungsprogramme entscheiden über die Verteilung von Ressourcen nach Kosten-Nutzen-Analysen, um Auswirkungen verschiedener Programme und Investitionen zu vergleichen. *Bewerten* und *Gewichten* sind somit zentrale Praktiken dieser komplexen Strategieentscheidungen, worauf auch die Analysten selbst verweisen: „Health policymakers across the globe are facing difficult financing decisions having to balance a large unmet and rising demand for health services, costly new drugs and technologies, ambitious international guidelines, and severely constrained health budgets“ (Remme et al., 2017). Insbesondere bei Infektionskrankheiten wie Malaria oder

[47] Beispiele sind der Streit um die Benzol-Grenzwerte in den USA in den 1980er Jahren und der folgende Einzug der Risikoanalyse mit einer Trennung von Risikobewertung und -management in der Regulierungspraxis.

Human Immunodeficiency Virus (HIV) im Globalen Süden, die einerseits massive wirtschaftliche Folgen haben und denen man andererseits oftmals in einem völlig unterfinanzierten und strukturell zerrütteten Gesundheitssystem begegnet, spielen nicht nur die erwarteten Investitionskosten, sondern auch die Abschätzung und Miteinrechnung der langfristigen finanziellen Konsequenzen eine entscheidende Rolle (Clinton & Sridhar, 2017; Stein & Sridhar, 2018).

Wie ökonomisierte Entscheidungsprozesse und insbesondere die Kosten Nutzen-Analysen in der internationalen Gesundheitspolitik zum Tragen kommen, lässt sich deshalb gut mit Beispielen der globalen *Roll-Back-Malaria*-Initiative (RBM) illustrieren (Kramer et al., 2009; RBM Partnership to End Malaria, 2015b; Steketee & Campbell, 2010). Dieses 1998 gegründete Netzwerk unter Federführung der Weltgesundheitsorganisation (WHO) versucht, die Anstrengungen von Organisationen der Vereinten Nationen, von Nichtregierungsorganisationen (NGOs), Stiftungen, Forschungs- und Entwicklungshilfeprogrammen sowie privatwirtschaftlichen und lokalen Initiativen im Umfeld der Malaria-Bekämpfung zu koordinieren und mit Ressourcen auszustatten (Fraser et al., 2006; Nabarro & Tayler, 1998). Hauptaufgabe der *Roll-Back-Malaria*-Zentrale in Genf ist es, Konsens über umsetzbare Strategien zwischen den über 500 beteiligten Organisationen herzustellen. Sie steht damit nicht nur vor der Herausforderung, eine globale Strategie zu entwickeln, die sich in diversen lokalen Kontexten, insbesondere in verarmten Regionen des südlichen Afrikas sowie Südostasiens, bewährt. Ziel von *Roll Back Malaria* ist ebenfalls, diese Strategieentscheidungen zwischen den Beteiligten zu kommunizieren und zu vermitteln. Neben einer Vielzahl von beteiligten Gesundheits- und Entwicklungshilfe-Organisationen müssen die betroffenen Länder und Regierungen von internationalen Strategien und gemeinsamen Zielvereinbarungen ebenso überzeugt werden wie die Geldgeber der Initiativen, die staatliche wie auch privatwirtschaftliche Stellen umfassen (Clinton & Sridhar, 2017; RBM Partnership to End Malaria, 2018b).

Entscheidungen müssen folglich in einem komplexen System getroffen werden, das auf international wie lokal tätige Akteure und ihre verschiedensten professionellen Hintergründe, Ressourcen, Motive und Interessen ausgerichtet ist. Wie im breiteren Feld der globalen Entwicklungs- und Gesundheitspolitik sind auch in der *Roll Back Malaria Partnership* Strategievorgaben durch formalisierte Entscheidungsprozesse zu

verzeichnen. Standardindikatoren und vergleichbare Bewertungsmaßstäbe dienen hier der Transparenz und Nachvollziehbarkeit von Zielvereinbarungen über fachliche und lokale Grenzen hinweg (Remme et al., 2001)[48]. Gleichzeitig bringt die Vielzahl der beteiligten Institutionen und Akteure einen gewissen Zwang zur Standardisierung und Ökonomisierung von Entscheidungsprozessen mit sich. Dieses Vorgehen hat erstens zur Folge, dass die möglichen Konsequenzen von *Roll-Back-Malaria*-Initiativen stets vom lokalen Kontext abstrahiert und kontextübergreifend systematisiert werden. Es werden also beispielsweise Infektions- oder Sterberaten in bestimmten Altersgruppen verglichen und mit der Einführung von bestimmten gesundheitspolitischen Maßnahmen ins Verhältnis gesetzt (RBM Partnership to End Malaria, 2015b, 2015c, 2017, 2018a). Zweitens bedeutet es, dass – im Sinne einer Kosten-Nutzen-Analyse – sowohl die aufzuwendenden Ressourcen wie auch die Effekte der *Roll-Back-Malaria*-Maßnahmen monetär bewertet werden. Wie bereits erläutert, zeigt sich auch hier das Problem des Einbezugs von nicht monetären Gütern, wie zum Beispiel der Gesundheit von Menschen, sowie der Abschätzung von unsicheren oder indirekten Langzeitfolgen beispielsweise durch Insektizideinsätze im Rahmen der Malariamückenbekämpfung (Goodman, Coleman & Mills, 1999; Guimarães et al., 2007; Kabasenche & Skinner, 2014).

Ein weiteres Problem evidenzbasierter Politik im Globalen Süden ist, dass sie mit Parametern und Determinanten operiert, die für deutlich reichere Länder entwickelt wurden[49]. Die Ökonomen Lisa Robinson, James Hammitt und Lucy O'Keeffe haben dies am Beispiel der VSL-Schätzungen *(Value per Statistical Life / Value of Mortality Risk Reductions)* erläutert, die breit für Entwicklungsprogramme eingesetzt werden, sich aber ursprünglich auf die US-amerikanische Wirtschaft und Gesellschaft bezogen bzw. für OECD-Mitgliedsstaaten entwickelt wurden. Allein die deutlich höhere Säuglingssterblichkeit, aber auch die generell geringere Lebenserwartung in Entwicklungsländern stellen Herausforderungen an die ökonomischen Analysen dar, denen in der Regel aber gar nicht begegnet wird. Robinson bilanziert beispielsweise bezogen auf die *Value-per-Statistical-Life*-Bewertung von Kindern und Erwachsenen,

[48] Bspw. Percentage reduction in mortality of children <5 years old, Percentage reduction in malaria incidence rate, Percentage of districts systematically using health information for planning etc.

[49] Eine der VSL-Schätzungen vergleichbare Kritik existiert für das Konzept der DALYs (*Disability-Adjusted Life Years* bzw. *Disease-Adjusted Life Years*) in globalen Gesundheitsprogrammen (Parks, 2014).

dass diese sich nach reichen Ländern richtet, weil dies die einzige Forschungsgrundlage ist. „For low- and middle-income countries, little empirical research is available and it is unclear whether the same patterns hold“ (Robinson, Hammitt, & O'Keeffe, 2019). Dieser von Robinson und anderen kritisierte Mangel an Forschung ist aber für Kollaborationen wie *Roll Back Malaria* kein Grund, von *Value-per-Statistical-Life*-Schätzungen abzusehen und vorerst weitere, kontextbezogene Forschungen abzuwarten. Vielmehr ist in den Publikationen der *Roll Back Malaria Partnership* von der Diskrepanz zwischen Entwicklungs- und Anwendungskontext oder einem folglichen Mangel an Evidenz keine Rede. Zu sehr steht dazu die Evidenzbasiertheit der Kampagnen im Vordergrund (Robinson, Hammitt, Jamison, et al., 2019; Yamey, 2001).

In den Publikationen der *Roll Back Malaria Partnership* sind Strategien und Zielvereinbarungen stets sehr allgemein formuliert, da sie auf internationales Publikum und in der Regel auch auf eine globale Anwendbarkeit ausgerichtet sind. Gleichzeitig spielen Verweise auf fortlaufende Evaluierung der eigenen Arbeit und Schlussfolgerungen aus dem Geleisteten eine entscheidende Rolle (RBM Partnership to End Malaria, 2015a, 2017, 2018a). Die Publikationen leben von einer Mischung aus Slogans wie beispielsweise „Vision: A world free from the burden of Malaria“ und sachlich aufbereiteten Textinformationen sowie Bildern, die Assoziationen bedienen sollen. Sie formulieren ihre Strategien kurz und einprägsam und kombinieren emotionalisierende Fotos mit Statistiken und Diagrammen. Insbesondere Fotos lächelnder schwarzer Kinder zeigen dem Leser, dass es hier um mehr geht als bloße Finanzkalkulationen, nämlich um die Entwicklung und um die Zukunft des afrikanischen Kontinents (RBM Partnership to End Malaria, 2017, 2018a). Dass ökonomische Analysen allgemeinpolitisch eingebettet und mit moralischen Werten versehen werden, findet sich auch in der *Roll Back Malaria* zugrundeliegenden Forschungsliteratur. „The compelling economic case for fighting malaria underscores the social and ethical merits of eradicating this disease“, schreibt beispielsweise die Forschungsgruppe um den Ökonomen Marc Purdy. „We hope that greater awareness of the economic case for eradication will better serve the humanitarian imperative of ridding the world of malaria“ (Purdy et al., 2013). Dieses Changieren zwischen zweckrationalen bzw. ökonomischen Gründen und dem Bezug auf humanitäre Ziele ist auch für die Frage nach Evidenzpraktiken interessant: Auf die politisch oder moralisch geframte Frage nach dem

Umgang mit sozialer Ungleichheit und unterschiedlichen Lebenschancen liefern hier ökonomische Berechnungen die Evidenz.

Wie groß dabei die Diskrepanz zwischen globaler Plakativität und konkreter Evidenz sein kann, zeigt beispielsweise eine der wichtigsten *Roll-Back-Malaria*-Strategien der letzten Jahre: die Einführung von mit Insektiziden behandelten Netzen (ITNs/LLITNs: *Insecticide-Treated Nets / Long Lasting Insecticide-Treated Nets*) zum Schutz vor dem Vektor der Krankheit, der Anopheles-Mücke. Diese Netze gehören seit den 1990er Jahren zum Instrumentarium der weltweiten Malariakontrolle. Dank ihrer ressourcenschonenden Anforderungen entwickelten sie sich rasch zu einer der Schlüsseltechnologien, die symbolisch für die Prinzipien globaler Gesundheitspolitik stehen: günstig in der Produktion, simpel in der Anwendung, global einsetzbar (Dolan et al., 2019; Hill et al., 2006; Webb, 2014). So vermerkt der *Roll-Back-Malaria*-Jahresbericht von 2017 die Verteilung von 582 Millionen *Insecticide-Treated Nets* zwischen 2014 und 2016 und unterstreicht den effizienten Charakter dieser Intervention (RBM Partnership to End Malaria, 2017).

Betrachtet man jedoch die Anfangsgeschichte und die fortschreitende Einführung der Netze, zeichnet sich dies keineswegs durch transparente und nachvollziehbare Bewertungs- und Gewichtungsprozesse aus[50]. Insektizidnetze wurden zunächst in den 1990er Jahren mit randomisierten kontrollierten Studien (RCTs) in Siaya in Westkenia erprobt. Die Historikerin Kirsten Moore-Sheeley beschreibt, wie Forscher der U.S. Centers for Disease Control und des Kenya Medical Research Institute kontinuierlich ihre Forschungsfragen und -praktiken der lokalen Bevölkerung und den Gegebenheiten anpassten, diese lokale Spezifik aber nicht in ihren Schlussfolgerungen erwähnten, geschweige denn die Reichweite ihrer Ergebnisse entsprechend eingrenzten (Krezanoski, 2016; Moore-Sheeley, 2017). In Anbetracht dieses Entstehungskontextes überrascht es nicht, dass die Einführung der Insektizidnetze dieselbe Kritik trifft, die auch gegenüber weltweiter Gesundheitspolitik im Allgemeinen geäußert wird: Globale Gesundheitstechnologien und -politiken ignorieren vielfach lokale Spezifika. Sie versäumen es, lokales Wissen zu nutzen und sie versuchen, Technologien zu

[50] Zur Diskussion von ITNs s. (Brieger, 2017; R. K. D. Peterson et al., 2011)

implementieren, die für andere Kontexte entwickelt wurden und nicht ubiquitär funktionieren (Adams, 2016; Barnes, Amy Parkhurst, 2014; Kouyaté et al., 2007).

Folglich zeigt sich auch in der lokalen Implementierung der *Roll-Back-Malaria*-Strategien, dass den formalisierten Kriterien der internationalen Analysen hier weniger Relevanz zukommt und sie keineswegs gleichmäßig angewandt werden. Beispielsweise haben der Zugang zu Medikamenten und weiteren Ressourcen sowie politische Machtverhältnisse und kulturelle Prägungen entscheidenden Einfluss auf den Erfolg internationaler Vorgaben (Pfeiffer & Nichter, 2008; Yamey, 2001). Im Versuch, lokale Spezifika zu berücksichtigen und bestimmte Zielgruppen zu erreichen, wenden lokale Akteure die Vorgaben aus Genf nicht nur äußerst freihändig an, sondern definieren auch eigene Kriterien für Erfolg und Misserfolg (Teklehaimanot & Mejia, 2008). Diese Diskrepanz zwischen lokalen und internationalen Bewertungskriterien unterläuft jedoch nicht nur die versprochene Transparenz der *Roll-Back-Malaria*-Initiative, sondern geht vielfach auch mit einer auf internationaler Ebene geringen Aufmerksamkeit für die tatsächlichen lokalen Ergebnisse und Möglichkeiten einher. In dieser Hinsicht lassen sich die Grenzen evidenzbasierter globaler Politik durchaus mit der im Kapitel *De- und Rekontextualisieren* referierten Kritik an evidenzbasierter Medizin vergleichen: Auch hier geht das Streben nach Evidenz und Vergleichbarkeit im Zweifel zu Lasten der im Einzelfall erzielten Ergebnisse.[51]

Das Problem der *Roll Back Malaria*-Initiative ist in dieser Hinsicht also nicht in erster Linie der grundsätzliche Mangel an Evidenz, sondern die fehlende Berücksichtigung ihrer Spezifik und deren Bewertung und Gewichtung. Bei der Einführung von *Insecticide-Treated Nets* gab es nach jahrelangen Experimenten durchaus Evidenz für ihre Wirksamkeit, nur war diese eben kontextgebunden. Diese Einschränkung ging jedoch in den ökonomisierten Entscheidungsprozessen unter. Am Beispiel der *Insecticide-Treated Nets* wird deutlich, dass evidenzbasierte Gesundheitspolitik nicht automatisch mit transparenten Entscheidungsprozessen einhergehen muss. Vielmehr wird, wenn die Gewichtung spezifischer Evidenz nicht offengelegt wird, damit der gesamte Entscheidungsprozess verschleiert. Bewertungskriterien für erfolgreiche *Roll-Back-*

[51] S. Kapitel *De- und Rekontextualisieren* in diesem Band.

Malaria-Maßnahmen müssten also viel stärker auf ihre Spezifik befragt werden und diese Spezifik sollte in die *Bewertung* und *Gewichtung* von Forschungsbefunden und Ergebnissen einbezogen werden.

Zweitens illustrieren Analysen im Zuge der *Roll-Back-Malaria*-Initiative ein weiteres generelles Problem evidenzbasierter Entscheidungsprozesse in verarmten Regionen der Welt: Diese Analysen basieren auf Parametern, die für einen anderen Kontext entwickelt wurden, nämlich für die reicheren Länder des Nordens. Versuche, Forschungen voranzutreiben, die Determinanten für *Low Income / Lower Middle Income Countries* entwickeln, stecken dagegen noch in den Anfängen. Hier zeigt sich also, was passieren kann, wenn die Forderung nach evidenzbasierter Politik auf eine Forschungslücke trifft: Anstatt diese Lücke zu markieren, wird der Forderung nach Evidenz nachgekommen. Inwieweit diese Evidenz aber zu den Anwendungskontexten passt, ist kaum nachvollziehbar.

Dieser Befund knüpft an die breitere These einer Ignoranz der (Entwicklungs-)Ökonomen gegenüber der Realität des Globalen Südens an, obwohl diese Experten zu genau diesen Regionen forschen und damit politische Entscheidungen beeinflussen. Abhijit Banerjee und Esther Duflo, die Autoren des des Grundsatzwerks *Poor Economics*, sehen darin einen entscheidenden Grund für die Ineffizienz der Entwicklungshilfe.

Wenn wir das träge, schematische Denken aufgeben, das jedes Problem auf die gleichen allgemeinen Prinzipien reduziert, wenn wir den Armen richtig zuhören und uns bemühen, die Logik ihrer Entscheidungen zu verstehen, wenn wir akzeptieren, dass wir uns irren können, und jede scheinbar noch so vernünftige Idee empirischen Tests unterziehen, dann werden wir nicht nur in der Lage sein, effektive Maßnahmen zu entwickeln, sondern auch besser verstehen, warum die Armen so leben, wie sie leben (Banerjee et al., 2012).

Liest man diesen Aufruf im Lichte unserer Überlegungen zu Evidenzpraktiken als Aushandlungsprozesse um die Gültigkeit von Wissen, wird klar, dass er nicht nur auf eine andere Politik, sondern insbesondere auf andere Modi des Überzeugens abzielt. Die Parameter, nach welchen wir Strategien der Entwicklungspolitik beurteilen, nach denen Expertise uns glaubwürdig scheint und Wissen als gesichert gilt, sind zu hinterfragen.

3.4.6. **Evidenz als Entscheidungshilfe in der Kritik**

Die Kosten-Nutzen-Analyse soll vielen Ansprüchen gerecht werden. Eine differenzierte Betrachtung von Kontroversen ermöglicht eine kritische Einordnung ihrer Ergebnisse und der Evidenzpraktiken *Bewerten* und *Gewichten* im Allgemeinen. Das Beispiel der *Roll-Back-Malaria*-Strategien illustriert vorrangig spezifische Probleme evidenzbasierter Politik im Globalen Süden. Die *Bewertung* und *Gewichtung* von Evidenz im Rahmen der Kosten-Nutzen-Analysen öffnet jedoch auch ein allgemeineres, fundamentaleres Feld für Kontroversen. Einige dieser Kontroversen sollen im Folgenden diskutiert werden. Sie entstehen aus mehreren Faktoren, die bereits im bisherigen Verlauf des Kapitels aufgezeigt wurden: Zunächst ist die Kosten-Nutzen-Analyse an einem sehr kritischen Punkt im Prozess der Konfliktbewältigung angesiedelt, weiterhin sind an ihre Ergebnisse weitreichende Entscheidungen geknüpft. Oft finden sie in komplexen Problemfeldern Anwendung, in denen eine Entscheidung über „akzeptabel“ oder „inakzeptabel“ keineswegs einfach möglich ist. Auch ist zu bedenken, dass sich Kosten-Nutzen-Analysen in einem westlichen und neoklassizistischen Kontext entwickelten, dessen Annahmen nicht problemlos auf andere Kontexte übertragen werden können.

Die erste Kontroverse entsteht aus der kritischen Verortung der Kosten-Nutzen-Analyse und bezieht sich auf die Verwissenschaftlichung von Diskursen und deren Auswirkungen auf demokratische Prozesse. Kosten-Nutzen-Analysen zielen eigentlich darauf ab, die wertenden Elemente im Prozess der evidenzbasierten Politik (vor allem im Hinblick auf effiziente Ressourcennutzung) zu „rationalisieren“. Gleichzeitig trägt diese Entwicklung aber auch zu einer Verwissenschaftlichung und Entrückung des Bewertungsprozesses bei, was eine Beteiligung der Öffentlichkeit erschwert. Die Voraussetzung technischer und wissenschaftlicher Analysen für eine Bewertung komplexer Situationen schreibt Experten automatisch eine hohe politische Autorität zu und erschwert Laien den Zugang zum Prozess, da sie nicht über das notwendige Wissen und/oder die entsprechenden Ressourcen verfügen (Kinchy, 2010). Für ein Verständnis des Zulassungsprozesses (und damit die Möglichkeit zur Kritik an diesem) von Pestiziden ist zum Beispiel umfangreiches naturwissenschaftliches Wissen aus der Toxikologie oder Ökotoxikologie notwendig. Diese Verwissenschaftlichung kann zu Anpassungsstrategien anderer Akteure führen, die ihrerseits ebenfalls (wissenschaftlich basierte) Instrumente und Strategien entwickeln, um sich am technischen Prozess zu beteiligen. Kinchy konnte zum

Beispiel in einer Studie zum Streit um die Zulassung von gentechnisch veränderten Maissorten nachweisen, dass Interessengruppen selbst begannen, wissenschaftliches Wissen zu nutzen und zu produzieren (Kinchy, 2010). Dies kann wiederum als problematisch angesehen werden, da sich hier besonders starke oder durch Geld unterstützte Gruppen durchsetzen können, die die Präferenzen der Gesellschaft möglicherweise verzerrt abbilden (Sunstein, 2018). Somit kann die Kosten-Nutzen-Analyse entweder als antidemokratisches oder als demokratiestützendes Instrument gesehen werden: Einerseits schließt sie durch Technisierung Laien aus und birgt die Gefahr der Verzerrung, schützt aber Entscheider und Bürger vor dem Einfluss von Lobbyismus, Ideologie oder öffentlicher Hysterie (Dobrow et al., 2004; Driesen, 2004).

Die zweite Kontroverse, die hier thematisiert werden soll, entsteht aus der Komplexität der Problemfelder und bezieht sich auf das Framing- und Wissensproblem der Kosten-Nutzen-Analysen. In dieser Kontroverse finden sich Argumente um die Methodik der Kosten-Nutzen-Analyse selbst. Kosten-Nutzen-Analysen funktionieren an sich als eine Art von Framing, das alternative Arten der Bewertung von Konsequenzen ausschließt (Saltelli & Giampietro, 2017). In diesen Analysen ist außerdem selten universell definiert, welche Arten von Kosten und Nutzen eingeschlossen werden (Driesen, 2004). Dadurch ist ein Framing, also der Ein- und Ausschluss von Effekten, abhängig von den durchführenden Experten. Effekte können durch Ignoranz externalisiert werden und finden dadurch keinen Eingang in den Entscheidungsprozess.

Diese Kritik ist übertragbar aus dem Diskurs Realismus – Konstruktivismus in der Risikoanalyse. Während im Realismus davon ausgegangen wird, dass die wissenschaftliche Bewertung die objektive Realität am besten abbildet, argumentieren Konstruktivisten, dass das Framing der Risikoanalyse nur Konventionen einer bestimmten Elite wiedergibt, indem eine Gruppe einen für die Gruppe logischen Konzeptrahmen bestimmt (Klinke & Renn, 2002). Vergleichbar argumentiert der Ökonom Friedrich von Hayek 1974 in seiner Analyse *Anmaßung von Wissen*, dass die „Planer“ nie über das gesamte, relevante Wissen aller Akteure verfügen können[52]. Dies hat, so schreibt Sunstein, weitreichende Folgen für die Rechtfertigung von Kosten-

[52] Hayek geht davon aus, dass sich dies durch Marktautomatismen lösen lässt. Dagegen sprechen Konstruktivisten im Risikodiskurs eher von Beteiligung der Gesellschaft (Hayek, 1945).

Nutzen-Analysen: „If cost-benefit analysis is essential to sensible judgements, incomplete knowledge, when it exists, would appear to be a serious and potentially devastating problem.“ (Sunstein, 2018, S.80). Dass unvollständige Wissensbestände aber häufig sind, ohne dass daraus Konsequenzen für globale Politik gezogen werden, hat das Beispiel der *Roll-Back-Malaria*-Analysen gezeigt. Auch in der Zulassung von Pestiziden macht die Komplexität der Folgen für Gesellschaft und Umwelt vollständiges Wissen zu einer Utopie: Wie bereits erläutert, ziehen Entscheidungen unterschiedlichste Effekte auf lokaler und globaler Ebene nach sich, meist verknüpft mit hoher Unsicherheit (Sexton et al., 2007).

Weiterhin bemängeln Kritiker von Kosten-Nutzen-Analysen, dass häufig die Verteilung von Nutzen und Kosten nicht ausreichend berücksichtigt wird (Fraiberg & Trebilcock, 1998; Sunstein, 2018). Dass dies in der Pestizidregulierung ein relevantes Kriterium ist, konnten die Agrarökonomen Lichtenberg, Parker und Zilberman bereits 1988 zeigen: Sie wiesen nach, dass sich die Wohlfahrtseffekte für verschiedene Akteure abhängig von den Angebots- und Nachfragereaktionen deutlich unterschieden. Potentielle Risk-Trade-Offs, welche im vorangehenden Verlauf des Kapitels erläutert wurden, werden somit vernachlässigt oder verschleiert. Ein weiterer Kritikpunkt ist, dass die Festlegung einer Rate zur Diskontierung von zukünftigen Risiken und Nutzen, z.B. bei generationsübergreifenden Effekten, schwierig ist und oft vernachlässigt wird. Auch Substitutions- und Mitnahmeeffekte sind Faktoren, die relevant sind, aber selten berücksichtigt werden.(Arrow et al., 1996; Fraiberg & Trebilcock, 1998)

Methodische Probleme betreffen insbesondere die Möglichkeit der Messung von Wohlfahrtseffekten im Allgemeinen[53]; und das Problem der Abbildung von Nutzen durch vermiedene Schäden im Besonderen. Bei der Berechnung des Nutzens von Pestiziden wird beispielsweise oft vernachlässigt, dass sie Qualitätsverluste durch Schadorganismen am Getreide verhindern können, obwohl diese ca. 20 Prozent des durch Pestizideinsatz erzielten Mehrumsatzes ausmachen. Verschiedene Studien zeigen, dass die Vernachlässigung der Qualitätseffekte zu einer Unterschätzung der Wohlfahrtseffekte führt (Babcock et al., 1992; Kawasaki & Lichtenberg, 2015). Weiterhin können die

[53] Eine Kosten-Nutzen-Analyse kann im schlechtesten Fall entkoppelt von Wohlfahrtseffekten sein, da sie nur ein Proxy für Wohlfahrt ist (Sunstein, 2018).

Messungen von Wohlfahrtseffekten, wie alle wissenschaftliche Studien, widersprüchliche Ergebnisse produzieren. So zeigen verschiedene Studien zum Einfluss vom Rückstandshöchstgehalten von Pestiziden auf den Handel mit den regulierten Lebensmitteln positive oder negative Effekte (Handford et al., 2015).

Die Messung von Wohlfahrtseffekten basiert weiterhin in vielen Fällen auf der vorangestellten Risikobewertung: Gesundheitseffekte von Pestiziden können erst dann monetär bewertet werden, wenn sie naturwissenschaftlich identifiziert und quantifiziert wurden. Erst wenn leberschädigende Eigenschaften eines Pestizids beziffert wurden, kann dieser Effekt monetär bewertet werden. Die Kosten-Nutzen-Analyse ist dadurch stark abhängig von der quantitativen Risikobewertung und übernimmt deren Problematiken, zum Beispiel in der Extrapolation von Effekten oder Datenlücken und den damit verbundenen Unsicherheiten, welche auch im Kapitel *Messen und Ermessen* dieses Bandes diskutiert werden. Zusammengefasst können Kosten-Nutzen-Analysen ein Framing- oder Wissensproblem haben, da sie sehr komplex sind, sehr viele Daten benötigen und auf vielen Annahmen und vorausgegangenen Verfahren beruhen.

Eine weitere Kontroverse im Kontext von Kosten-Nutzen-Analysen betrifft moralische Fragestellungen. Dazu gehört zunächst die implizierte Notwendigkeit der Bewertung und Quantifizierung von menschlichem Leben oder anderen nicht monetären Effekten. Dies macht fragwürdige Annahmen über den Wert des menschlichen Lebens notwendig, für die es bisher keine unumstrittene Methode gibt (Driesen, 2004), auch wenn dieses Problem in der Wissenschaft bereits seit Starr (1969) in verschiedensten Ansätzen behandelt wird (Starr, 1969). Dies wurde bereits am Beispiel der Pestizide und der *Roll-Back-Malaria*-Initiative im vorangegangenen Verlauf des Kapitels deutlich: Zum einen können unterschiedliche Wertzuschreibungen für die Leben von verschiedenen Akteuren kritisch betrachtet werden (Konsumenten vs. Anwender). Zum anderen kann die Übertragung des *Value per Statistical Life* in unterschiedliche Kontexte problematisch sein[54]. Ähnlich verhält es sich mit der Bewertung von anderen nicht monetären Konsequenzen wie zum Beispiel Effekten auf Biodiversität oder der Bewertung von Zufriedenheit oder moralischen Bedenken, deren direkte Wohlfahrtseffekte für eine

[54] Man spricht hier auch von *benefit transfers* (Übertragung von Nutzen). Die in (kontingenten) Bewertungen gemessenen Werte für nicht monetäre Güter werden auf einen anderen Kontext übertragen.

Gesellschaft oft schwierig zu erkennen und quantifizieren sind.(Fraiberg & Trebilcock, 1998) Auch stellt sich die Frage, wie mit den Ergebnissen von Kosten-Nutzen-Analysen verfahren werden soll, wenn deren Ergebnisse gesellschaftlichen Konventionen widersprechen oder moralische Grenzen überschreiten.[55] Dieses Problem moralischer Aufladung sachlicher Analyse zeigt sich insbesondere in Feldern, die nicht unabhängig von größeren politischen Problemlagen zu lösen sind. Die weltweite Bekämpfung von Krankheiten wie HIV oder Malaria etwa lässt sich nicht unabhängig von Fragen nach Gerechtigkeit und Verteilungskonflikten betrachten, so dass auch die dazugehörigen ökonomischen Berechnungen immer einen Kommentar oder sogar eine Handlungsempfehlung zum Problem der globalen Armut beinhalten.

3.4.7. Schlussfolgerungen

Dieses Kapitel zeigt, welche Bedeutung die Praktiken *Bewerten* und *Gewichten* im Feld der internationalen Gesundheits- und Umweltpolitik haben und welche Möglichkeiten und gleichzeitig Konflikte und Kontroversen diese zentrale Stellung der Kosten-Nutzen-Analyse eröffnet. Indem Entscheidungsprozesse in komplexen Problemlagen mit Kosten-Nutzen-Analysen legitimiert werden, haben sie sich nach und nach als Schlüsselverfahren für *Bewertungen* und *Gewichtungen* in der evidenzbasierten Politik etabliert und damit alternative Bewertungsverfahren verdrängt. Sie dienen als Möglichkeit, diverse Interessen und Effekte durch Monetarisierung auf eine vergleichbare, argumentative Ebene zu stellen. Die Verwendung etablierter ökonomischer Verfahren ermöglicht idealerweise, den Ansprüchen an Bewertungs- und Gewichtungsprozesse in Bezug auf intersubjektive Nachvollziehbarkeit, Transparenz, Wiederholbarkeit und Glaubwürdigkeit gerecht zu werden. In den Beispielen des Kapitels zeigt sich, dass dies keinesfalls eine einfache Aufgabe ist. Die zunehmende Formalisierung des Verfahrens führt automatisch zum Ausschluss von bestimmten Faktoren – der Anspruch an die Replizierbarkeit von Entscheidungen impliziert gleichzeitig einen Verlust der Anpassungsfähigkeit. Die Beispiele dieses Kapitels zeigen, wie anspruchsvoll die Forderung nach Evidenz als gemeinsame Basis für die Begründung von Entscheidungen

[55] Sunstein nennt hier als Beispiel Tierwohl – auch wenn die Zahlungsbereitschaft niedrig wäre, ist es moralisch geboten, dass Tiere nicht misshandelt werden dürfen (Sunstein, 2018).

sein kann. Die verschränkte Analyse von Bewertungs- und Gewichtungsprozessen hat zudem das Potential, die Bedeutung impliziter Faktoren sichtbar zu machen.

Die Komplexität der geschilderten Entscheidungen in der internationalen Umwelt- und Gesundheitspolitik erfordert eine stetige Aushandlung, wie Risikofaktoren und Folgeabschätzungen in die Entscheidungsfindung einbezogen werden sollen. Die Frage, wie Daten in diesen Auseinandersetzungen beurteilt werden und dadurch wertbasierte Urteile rechtfertigen und begründen, hat dabei auch eine politische Dimension: Von Bewertungen und Gewichtungen hängen weitreichende politische Entscheidungen, die Verteilung von Ressourcen und die Ausrichtung großangelegter Programme ab. Die Evidenzpraktiken des *Bewertens* und *Gewichtens* sind damit beides: Einerseits Machtinstrumente der Politik, andererseits Instrumente im Modus des Begründens und der Konfliktbewältigung.

3.4.8. Acknowledgements

This article is based on research supported by funding from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Research Group 2448.

4. Conclusion

The following chapter provides a joined conclusion on the studies and results of this dissertation. The first section lays out main findings and discusses them. The second section is concerned with implications for future research while the third section focusses on implications for risk analysis.

4.1. Main findings and discussion

Food safety determination is a societal challenge because it affects various individual actors or groups in different ways and can have an impact on their health or welfare. It further interacts with other factors like environmental impact or food security and societal development. Thus, the determination of socially optimum food safety levels is of interest for the whole society. The discussion on food safety involves multiple disciplines which contribute different explanations, facets, and concepts. Additionally, drivers like globalization or moralization of food complicate value chain structures and consumer behavior. According to regulatory frameworks, food safety determination should be based on evidence which is challenging if evidence is understood as socially accepted knowledge. It has been proven to be problematic to assume that evidence-based risk assessments and safety standards are *per se* conflict- and value-free, most recently in the debate on the re-approval of the herbicide Glyphosate.

Additionally, consumers' safety determination does often not comply with experts' risk assessment conclusions. So far, it was not possible to solve these conflicts only by the provision of information through risk communication and to transform empirical evidence into socially accepted knowledge. Thus, it is crucial to understand consumers' underlying evidence determination processes.

The aim of this dissertation was to analyze processes of evidence determination for food safety and to investigate interactions between scientific and consumer public in these processes. Four studies have been conducted to address these aims from different perspectives: A conceptual perspective, a perspective on consumer food safety criteria, a perspective on the practice of risk assessment, the core instrument to determine evidence for food safety, and a perspective on evaluating and weighting evidence in cost-benefit analysis.

It focused on the following primary research questions: *How does the determination of sufficient evidence for food safety differs between the science-based risk analysis process and intuition-driven consumers? How do these evidence practices interact in the determination of food safety and influence the selection of evidence for food safety?*

Based on qualitative data from literature, expert interviews, and in-depth interviews, the studies provide a holistic view on food safety determination, focusing on the evidence practices of science-based risk analysis and consumers. It focusses on food safety determination in the European regulatory system.

The dissertation contributes to research on food safety determination, risk analysis, and risk perception. It aims to provide a new approach to conceptualize food safety determination, to uncover the related evidence practices, and thus to identify current strengths and weaknesses in regulatory practice as well as to establish a new perspective on consumers in the process. Additionally, it shows food safety determination as a societal process.

Study 1 (section 3.1.) provides a conceptualization of food safety determination and the related evidence practices. Hereby, it focusses on the practices of risk analysis and consumer determination. It aims to provide understanding of the complex phenomenon and to identify food safety criteria used by different disciplines and actors. Result prove that it is not possible to establish a universal definition for food safety. Instead, the study introduces a conceptual framework to describe food safety based on the trading zone concept. From both actors, food safety is determined independently on different levels in form of food safety criteria. Risk analysis includes knowledge criteria from risk assessment and value criteria from risk management which result in standards. Consumers determine knowledge criteria including the relevant dimension for safety and value criteria in form of valuing decisions about safety which result in behavior. So far, this diversity of criteria has not been systematically analyzed and structured. Besides criteria, the study identifies the related evidence practices and difficulties. Its findings confirm that evidence for safety is a societal construct rather than a single empirical value.

This conceptualization improves the understanding of food safety determination by different actors. It acknowledges food safety determination as a societal process, including contributions from different disciplines. It helps to identify inconsistencies in evidence determination, problems in regulatory processes, roots of misunderstandings (between societal groups or scientific disciplines) and explain phenomena like “lay” criticism on scientific assessments. The study might support the inclusion of a new perspective on

consumers in the process, acknowledging their evidence practice as overlapping, but independent from risk analysis. It points to the need for a re-consideration of the role of consumer evidence in food safety determination, especially the question which type of consumer food safety criteria are useful for regulatory decisions. Additionally, it highlights the importance of economic approaches to structure the determination of value criteria in risk management.

Study 2 (section 3.2.) deepens the understanding of consumer food safety criteria and aims to add empirical evidence on the conceptualization in study 1. It employs an inductive approach and aims to provide a holistic perspective on food safety criteria used by consumers beyond existing frames and concepts. The results prove two main types of consumer food safety criteria: Safety dimension criteria (what is the constitution of safe food) and safety evaluation criteria (what is used for the evaluation of safe food). For each category, it identifies a rich set of subcategories which provides a detailed description of the qualitative understanding of food safety by consumers. It confirms that it might be insufficient to use simple risk perception scales to describe food safety determination holistically.

This analysis confirms the conceptual categorization of study 1. The determination and understanding of safety by consumers are based on a complex set of criteria which is underestimated when using the term "intuition" for it. It is important to differentiate between criteria which are used in the framing of food safety (dimension criteria) and potentially biasing or anchoring elements (evaluation criteria). It is possible to identify and separate these criteria in qualitative analysis. That might be helpful for complex risk analysis settings or for optimizing integration processes of societal information in risk analysis.

Study 3 (section 3.3.) analyzes the practical impact of study 1. It is concerned with the evidence practice of risk assessment and its interactions with different stakeholders. Whilst focusing on consumers, it also includes food production as second stakeholder group to have a control case and to identify similarities and differences. The study includes risk communication, because it is an important tool for stakeholder interaction in risk assessment organizations. The purpose of the study is to show contact zones between different evidence practices and how these are implemented and impacts the strongly codified practice of risk assessment. Additionally, it provides insights in related difficulties for risk assessors. The study identifies a variety of one-way and two-way interactions between risk assessment organizations and stakeholders in different phases of evidence determination: selection, processing, and use of evidence. It shows that there are a variety of direct and indirect

interaction mechanisms which contribute to evidence determination but also cause conflicts and difficulties for risk assessors. Stakeholders can be passive (objects to investigate) or active (objects who contribute) elements in these interactions.

This approach will prove useful in expanding our understanding of evidence determination in food safety in risk assessment organizations. It acknowledges that members of such organizations act in a societal context and have contact zones with related stakeholders in the evidence determination process. The empirical findings in this study contribute to recent theoretical and conceptual research on the role and the importance of stakeholders in risk assessment. The study highlights the need for structured approaches to identify and integrate diverse stakeholder evidence in different phases of risk assessment. This might help to prevent members in risk assessment organizations from value conflicts and increase the social stability of risk assessment.

Study 4 (section 3.4.) elaborates on the category of value criteria in risk analysis which were identified in study 1. Cost-benefit analysis is an established procedure in risk management which uses a monitorization approach to make effects of different nature comparable. The study aims to describe and critically discuss cost-benefit analysis as decision tool to evaluate and weight evidence. It shows cost-benefit analysis as an evidence practice itself, including mechanisms, its way to an established decision tool, and the critical issues. It shows that, on the one hand, codified evidence-practices aim to provide intersubjective transparency and comparability and thus may function as tool for reasoning and conflict management. On the other hand, it is likely to have decreased flexibility, to be culturally biased, and to exclude certain factors. Thus, it may function as an instrument of power.

This approach contributes to a reflective view on cost-benefit analysis. It acknowledges that evaluating and weighting evidence in policy is an evolving evidence practice itself. It highlights the relevance and threats of such established evidence practices. The elaborations of this study contribute to conceptual and historical research on the role and the mechanisms of cost-benefit analysis. This might help to interpret and reflect policy decisions.

This dissertation provides deep, reflective insights into evidence determination processes for food safety. It contributes to a stream of literature which supports the view on risk analysis as an interwoven complex societal process which needs to establish a new perspective on stakeholders (Devos et al., 2019; Klinke & Renn, 2002; van Zwanenberg & Millstone, 2000; Vareman & Persson, 2010). It extends the perspective from “how safe is safe enough?” (Fischhoff et al., 1978) to “what is meant by safety”? If evidence is a base for solving societal

conflicts it is important that there is a common understanding and agreement on the meaning of evidence (Zachmann & Ehlers, 2019). Thus, the understanding of evidence practices and its contact zones is a crucial task.

4.1.1. The trading zone concept as research framework

This dissertation focusses on evidence practices based on the trading zone concept. Results support the usefulness of the trading zone concept and the perspective of independent but interacting evidence practices. The four studies provide a holistic view on the actors' network presented in section, study 1 in form of a conceptualization, study 2 in form of an analysis of consumer food safety criteria, study 3 in form of a closer look on contact zones between science and consumers, and study 4 in form of a reflective view on evidence-based policy. Food safety determination shows central characteristics described by Galion (2010): First, results from study 1 show that it is possible to describe evidence practices from risk analysis and consumers as co-existing and not hierarchical. The conceptualization identifies relevant safety dimensions for consumers which are not based on the results of risk analysis. That was confirmed in a follow-up study 2 on relevant safety dimensions for consumers which shows that these dimensions are not based on classic risk metrics but on fundamental values. A second characteristic of the trading zone concept is that meanings of certain terms and concepts are not shared within both groups (Galison, 2010). This is consistent with findings from the four studies. Study 1 shows this lack of shared meaning on the level of safety definition and differences in food safety criteria. Although similar terms are used, they develop in different contexts and evidence practices. Study 2 elaborates on that and shows that consumers have their own understanding of safety based on a complex set of criteria. Study 3 shows the practical effects of these differences in risk assessment. Although not the focus of the analysis, results indicate that misunderstandings between risk assessors and consumers might be caused by different meanings of, for example, public participation or safety. The focus on evidence practices offered insights in safety determination and roots of such differences. Study 4 shows the effects of these differences in risk management practice. The analysis proves that political decision-makers need to evaluate and weight different understandings of safety and find a comparable base or a common language – in this example the monetarization in cost-benefit analysis. The third characteristic in trading zone concepts is a power imbalance between two actors. The imbalance of power between the evidence practices is significant: Science-based risk assessment is a key tool in European food safety

determination. The risk analysis concept dominates the process of evidence determination in food safety. Although there is a lot of research on consumer safety determination, it does primarily aim for an explanation or prediction of "biased" risk perceptions and see consumers as passive element in this process (s. study 3).

Thus, the trading zone concept provides a suitable approach to describe the food safety determination process in a new way and to adopt a perspective which supports constructivist streams of literature in risk research (Klinke & Renn, 2002; van Zwanenberg & Millstone, 2000).

So far, the field of food safety determination is dominated by the view of risk perception as reaction to empirical knowledge from risk analysis and experts. It includes (1) research on individual characteristics and beliefs which influence these perceptions, (2) research on hazard characteristics which bias this perception, and (3) research on risk communication to address these differences in perception. This approach has not been entirely successful so far. The large amount of research on risk communication, which focusses on changing attitudes or behavior, did contribute to the knowledge which messages are effective in which groups (Frewer et al., 2016). Nevertheless, neither risk perception nor risk communication research was able to solve the gap between risk analysis outcomes and risk perception which can be seen at skepticism on chemicals or new food technologies like genetically modified organisms (Lusk et al., 2014).

The trading zone concept provides a different perspective on consumers as independent objects in food safety determination without contradicting standing concepts. It integrates existing approaches of risk perception research into different levels of food safety criteria.

In general, risk analysis is a standardized concept codified in various documents and regulations, for example in the EU General Food Law (EC & EP, 2002). Nevertheless, food safety determination is concerned with a variety of hazards which are related to different disciplines. Thus, the evidence practices are likely to differ between cases. To allow for extensive qualitative analysis, two cases with different characteristics have been chosen for this dissertation, PPPs and microbiological hazards. The cases have been sampled purposefully following a maximum variation approach (Patton, 2015). The results of the studies show that this differentiation was useful. Hazard characteristics has proven to be influential on the evidence determination process.

4.1.2. Reflection on Quality

Specific methodological limitations are discussed in the studies, for example the interpretation of expert knowledge. Thus, in this section, quality criteria are discussed for this dissertation in general. Quality criteria of qualitative research differ from quantitative research (Bitsch, 2005). Bitsch (2005) listed four relevant quality criteria and related indicators: credibility, transferability, dependability, and confirmability.

Credibility is described as equivalent to internal validity and is differentiated as follows: "[...] in a qualitative research context, correspondence with reality is replaced by correspondence of the perspectives of the participants with the description of their perspectives by the researcher" (Bitsch, 2005). In this dissertation, different tools were applied to ensure credibility: In the literature review, a systematic search strategy and snowball procedure was applied to avoid a limited focus on certain areas and confirmation bias. In the in-depth interview study, a purposeful sampling was applied which ensured a broad representation of gender, age, and income. In the expert interview study, a document analysis was conducted in advance to gain a broader picture of the field of interest. Additionally, in-depth interviews allowed experts to explain their views. In general, studies and findings were discussed in conferences and research retreats with scientists from different disciplines before publication to control for biased interpretations. A further argument for credibility is the use of triangulation of different qualitative methods to provide different perspectives on the same phenomenon (Bitsch, 2005). *Transferability* as equivalent to external validity "refers to the degree to which research results can be applied to a context apart from where they were gained or with different subjects" (Bitsch, 2005). In qualitative research, this criterion is not a concern of the conductor of a study. The applying researcher needs to decide if concept fits in his/her context. Therefore, indicators for transferability are thick descriptions and purposeful sampling (Bitsch, 2005). The studies of this dissertation provide detailed descriptions of cases and contexts and rely on purposeful sampling of most informative cases including reasons for choices. *Dependability* "refers to the stability of findings over time" (Bitsch, 2005). Qualitative methods are characterized by evolving concepts, constructs, and methods. As indicator, it is important to document these changes (Bitsch, 2005). *Confirmability*, in quantitative research described as objectivity " deals with the issue of bias and prejudices of the researcher" (Bitsch, 2005). It requires a documentation of research process and interpretations, which is provided in study documentations (Bitsch, 2005).

4.2. Implications for future research

The dissertation investigates a topic which is related to various disciplines. It includes natural sciences in risk assessments, policy science, food (safety) economics, consumer research, risk perception research, and risk research in general. Thus, it entails different implications for future research.

Study 1 establishes a framework to describe food safety criteria from different disciplines. It entails various implication for consumer research in the field of risk perception or safety determination, not primarily of methodological nature but rather for conceptualization and framing of studies. First, it differentiates between biasing or influencing elements on perception on the one hand and criteria for safety determination on the other. That supports the need to acknowledge food safety determination not as attitude but as a complex evidence practice which can be analyzed on different levels. Second, it has implications for the explanation of the attitude behavior gap. If behavior is seen as a food safety criterion evolving as result of attitudes, it might be a helpful approach to label the difference between attitude and behavior not as a gap but as a tolerance range, also depending on barriers to behave according to their attitude (Hirschman, 1971). Based on this argument, it might be important to research on thresholds for behavioral changes. Third, results show that the basis of consumer safety determination is not well researched if consumers are seen as independent from risk analysis. So far, most research concentrated on behavior, attitudes or influencing elements in perception but did not focus on the underlying relevant dimensions for safety determination. Even if studies focused on these dimensions they were framed towards their influence on risk perception not as the basis for risk evaluations (for example Slovic, 1987). The study indicates that it might be useful to identify these dimensions and to differentiate them from attitudes. Based on the assumption that these dimensions are independent from risk analysis it might be important to assess them qualitatively to get out of framings of other evidence practices like risk analysis. These implications are supported by three case-studies which show that it is possible to analyze underlying dimensions qualitatively and that these dimensions are different between different case studies and genders (Bieberstein & Roosen, 2015; Hassauer & Roosen, 2019). For more evidence on these underlying dimensions, it would be necessary to develop specific (mixed) methods and systematic data collection schemes. This approach might also be interesting for risk communication research. So far, risk communication research focused mostly on the change of attitudes or behavior (Frewer et al., 2016). Here, the differentiation of food safety criteria

used by consumers might be an interesting approach. It might help to understand the fundamental meanings of safety and address them in communication.

Generally, study 1 shows the value of conceptual research for risk research to analyze and structure complex, multidisciplinary phenomena and supports the usefulness of such approaches in addition to empirical research (Aven, 2018).

Study 2 provides a qualitative analysis and categorization of food safety criteria used by consumers. It highlights the need for a differentiation between safety dimension criteria and safety evaluation criteria. It offers an approach to analyze these criteria inductively beyond existing risk perception concepts and proves that consumers construct an independent understanding of safety. Additionally, it confirms that it might be useful in research to differentiate between biasing elements (evaluation criteria – risk perception) and the framing of safety (dimension criteria – safety determination). The analysis shows that it is possible to identify such complex sets of criteria in in-depth interviews. For future consumer research, it might be useful in understanding complex risk settings or in combination with quantitative research in mixed-method studies. Follow-up studies might focus on the relevance of criteria or the comparison of hazard-specific criteria.

Study 3 identifies interactions with stakeholders in the evidence determination process of risk assessment organizations. The study uses two food safety cases in a maximum variation strategy and was able to produce a literal replication. This indicates that the basic logic of interactions with stakeholders is likely to be applicable for other cases (Ridder, 2017). The study provides a first attempt to structure interactions with stakeholders in different stages of risk assessment. Future research might use more case-studies or quantitative approaches to back up the systematic with more empirical evidence. Additionally, it might be worth to investigate specific links between interaction mechanisms and their impact on risk assessments for example in qualitative field observations to address implicit stakeholder integration mechanisms and their effects.

As indicated by study 1, study 3 highlights the need to evaluate participation barriers for stakeholders, theoretically described by Hirschman (1971). This is especially important for stakeholders which use different evidence practices than risk analysis.

In a side note, study 3 indicates that differences in perception of the roles of risk assessment are relevant for consumers – the approving role in the case of chemicals, or the saving role

in the case of biological hazards. It might be helpful to differentiate between cases if trust in organizations is analyzed.

Study 4 describes mechanisms, evolvement, and critical issues of cost-benefit analysis as a tool in evidence-based policy. The study offers a literature-based reflective perspective. Future research might focus on the critical issues and investigate in empirical studies how they are handled in practice. Additionally, study 4 might provide approaches for a conceptual improvement of cost-benefit analysis.

4.3. Implications for risk analysis

The findings of this dissertation have several important implications for the conceptualization and practice in risk analysis.

Most importantly, this dissertation allows a new perspective on food safety determination as a mix of evidence practices in a complex societal process. Based on this perspective, it is not possible to define “safe food” based on one empirical value, even if it is evidence-based. Food safety is an interdisciplinary phenomenon, including different forms of evidence.

The BSE crises and its effects on the European food safety system proofed that it is possible that public pressure can change an established system if it is not able to provide acceptable food safety levels (Vos, 2000). The introduction of the risk analysis principle shifted European food safety policy towards more evidence-based regulations; independency of risk assessment; transparency; and the recognition of responsible, informed consumers (van der Meulen & van der Velde, 2004). Recently, a similar form of public pressure on the current system evolved in the discussion about the re-approval of Glyphosate (Morvillo, 2020). These discourses are closely connected to the definition of acceptable risk levels. This dissertation adds a new perspective on the roots of such discourses. It analyzes different evidence practices and thus different forms of food safety determination by science-based risk analysis and by consumers. So far, society managed to bring these different evidence practices together more or less successful. The analysis of evidence practices and their contact zones identifies critical points in food safety determination which entail implications for optimization potential in risk analysis. These implications are discussed in the following section, separated by risk assessment, risk management, and risk communication.

4.3.1. Risk assessment

First, results support the need to acknowledge the existence of societal influences on risk analysis, and more specifically on the practice of scientific risk assessment which is the central element in food safety evidence determination. This entails a necessity for reflection and documentation of these influences to establish trust in these organizations and to prevent them from conflicts (Devos et al., 2019). There have been first attempts to systematically describe different stakeholder engagement mechanisms in the literature such as Rowe & Frewer (2005). Building on this, study 3 provides a systematical reflection of stakeholder interactions in different stages of risk assessment as contact zones between two different evidence practices. The results of this study show that it is important to define processes and quality criteria also for non-scientific information which emerge from different evidence practices, for example from consumers.

Second, objectives in stakeholders participation in risk assessment can be diverse and can vary from the satisfaction of regulatory requirements to the role of the public as quality assurance (Homan et al., 2001). The recent amendments of the General Food law did not systematically change the role and objectives of stakeholder participation. Amendments target more openness of risk assessment and optimization of risk communication but not really inclusion of (alternative) consumer evidence or significant changes in risk assessment practice (Chatzopoulou et al., 2020). It might be worth to investigate and reflect the current objectives of stakeholder participation. Stakeholder participation can be a valuable tool in framing of risk assessments, add local knowledge, discuss alternatives, or a quality control mechanism (Homan et al., 2001; Klinke & Renn, 2002).

The current role solidifies the construct of hierarchical evidence practices in European food safety determination with a strong focus on evidence from the natural sciences. It is unquestionably important that food safety regulations are based on scientific evidence for various reasons. Stakeholder or consumer engagement has difficulties: The lack of resources, delaying of processes, or the variety or instability of stakeholder concerns (Barker et al., 2010; Gruev-Vintila & Rouquette, 2007). Participatory information should not replace scientific information but add additional forms of evidence (Barker et al., 2010). Therefore, it is important to reflect the form of stakeholder evidence which is used in different stages of risk analysis.

Based on the conceptual framework established in study 1 and the empirical evidence in study 2, it is possible to critically reflect the level of consumer evidence which is included in stakeholder participation in which stage of risk analysis. For example, the framework shows that many individual factors influence food safety criteria in perception. Thus, the measurement of concern might be unsuitable because it does not acknowledge the complexity of underlying dimensions. A qualitative analysis of relevant dimensions of safety might be important in the framing of risk assessment. Another example would be the use of behavior as safety criterion which is closely connected to the possibilities to behave. Thus, it is important to reflect barriers for stakeholders and their possibilities to express disagreement to interpret consumer behavior and to establish efficient ways how stakeholders can express dissatisfaction (Fischhoff, 1994; Hassauer & Roosen, 2020; Hirschman, 1971; Kinchy, 2010; Sunstein, 2018). So, an optimal integration of stakeholders does not only require evidence on their safety determination but also contextual information.

Critical consideration of the form of consumer evidence and stakeholder activity can also support risk assessors. They are expected to fulfil expectations of independency and scientific integrity while acting in a complex social environment and facing the ambiguities of risk assessment (Johansen & Rausand, 2015; Vareman & Persson, 2010). The identification and systematic integration might help to prevent risk assessment from the accusation of informal influences. Hereby, it has been shown that there is a fundamental conflict between maintaining independency of risk assessment and involving stakeholders which is important and necessary. It is important to acknowledge that also risk assessments might come to different conclusions and involve scientific uncertainty.

4.3.2. Risk management

Risk management is responsible for the determination of value criteria, the “acceptability” of risk. Thus, it is important to acknowledge these value criteria as part of the evidence practice for food safety determination.

Results of this dissertation point toward the importance to develop economic approaches to determine social optimum levels in food safety which consider scarce resources and effects on trade, individuals or producers (Henson & Traill, 1993). This importance is also indicated by the evolving relevance of the research field of food safety economics which aims to provide methods to monetarize effects for different actors (Focker & van der Fels-Klerx, 2020). More systematic integration of this form of evidence into food safety determination

is necessary. Economic evidence needs to be seen as part of the evidence practice of food safety regulations. Nevertheless, also these forms of evidence have its issues which are extensively discussed in Study 4. Additionally, economic approaches should not replace evidence from natural sciences but help to make informed decisions about acceptability of safety levels.

In this context, it is important to acknowledge the existence of socially optimum levels of safety which are not necessarily equivalent with thresholds from natural sciences considered as safe (Henson & Traill, 1993). This might implicate higher risk levels than proposed by scientific considerations in some cases. Individual optimum levels might differ between stakeholders. The invention of systematic approaches for analysis and weightings in form of socio-economic methods would be crucial to optimize the evidence practice for acceptable food safety levels. Nevertheless, the existence of different evidence practices makes it difficult to find a common argumentative level. Study 4 shows the example of monetarization in cost-benefit analysis.

Besides suitable methods, there remains the question who is responsible for risk evaluations and providing societal information. Since the invention of the General Food Law, scholars have critized this lack of systematic risk evaluation at the interface between risk assessment and risk management and developed approaches for solutions (for example König et al., 2010). The responsibility question concerns also the question of framing risk assessments as discussed in the previous section on risk assessment (Vareman & Persson, 2010). It is not possible to separate risk assessment and risk management completely (Robinson et al., 2016). Risk assessment needs to provide policy-relevant knowledge in efficient ways (Ruzante et al., 2010; Zwietering, 2015). Risk assessment and risk management are associated with similar values including extra-scientific ones (Vareman & Persson, 2010). The introduction of a risk evaluation stage would acknowledge this problematic and provide a linking element between both.

4.3.3. Risk communication

Although risk communication has not been the focus of this dissertation, there are some indirect implications for it.

Study 2 identifies a complex set of food safety criteria used by consumers. These criteria can be differentiated in framing elements (dimensions of safety) and potentially biasing elements

(evaluation criteria). This might also be useful for the understanding of risk perception and thus, optimized communication.

The differences between the evidence practices support the explanation of a communication vacuum described by Leiss and Powell (2004). The meanings of the same terms might not be shared or understood by different groups because they have a different understanding of safety. If this vacuum is not addressed in communication, it leaves room for other communicators like different interest groups which might contribute to amplification phenomena known from risk perception research.

It might be an interesting question if risk communication could be the provider of social or economic evidence discussed in the previous section.

The previous arguments prove the usefulness of a new perspective on the well-known phenomenon of food safety determination and risk perception. The structured analysis on the roots of food safety evidence might increase mutual understanding, explain differences between societal groups, and provide approaches for an optimization of risk analysis.

5. References

Adams, V. (2016). *Metrics: What Counts in Global Health*. Duke University Press. https://doi.org/10.1215/9780822374480

Ågerstrand, M., & Beronius, A. (2016). Weight of evidence evaluation and systematic review in EU chemical risk assessment: Foundation is laid but guidance is needed. *Environment International*, *92–93*, 590–596. https://doi.org/10.1016/j.envint.2015.10.008

Alhakami, a S., & Slovic, P. (1994). A psychological study of the inverse relationship between perceived risk and perceived benefit. *Risk Analysis : An Official Publication of the Society for Risk Analysis*, *14*(6), 1085–1096. https://doi.org/10.1111/j.1539-6924.1994.tb00080.x

Ansell, C. K., & Vogel, D. (2006). *What's the beef? : the contested governance of European food safety*. MIT Press. https://books.google.de/books/about/What_s_the_Beef.html?id=BSNd7OrVv9gC&redir_esc=y

Arrow, K. J., Eads, G. C., Hahn, R. W., Lave, L. B., Noll, R. G., Portney, P. R., Russell, M., Schmalensee, R., Smith, V. K., & Stavins, R. N. (1996). Is there a role for benefit-cost analysis in environmental, health, and safety regulation? *Science, New Series*, *2*(2), 195–221. https://doi.org/10.1017/S1355770X97220164

Aven, T. (2018). Reflections on the Use of Conceptual Research in Risk Analysis. *Risk Analysis*, *38*(11), 2415–2423. https://doi.org/10.1111/risa.13139

Babcock, B. A., Lichtenberg, E., & Zilberman, D. (1992). Impact of Damage Control and Quality of Output: Estimating Pest Control Effectiveness. *American Journal of Agricultural Economics*, *74*(1), 163. https://doi.org/10.2307/1243000

Banerjee, A. V., Duflo, & Warmuth, S. (2012). *Poor economics: Plädoyer für ein neues Verständnis von Armut*. btb Verlag.

Barker, G. C., Bayley, C., Cassidy, A., French, S., Hart, A., Malakar, P. K., Maule, J., Petkov, M., & Shepherd, R. (2010). Can a participatory approach contribute to food chain risk analysis? *Risk Analysis*, *30*(5), 766–781. https://doi.org/10.1111/j.1539-6924.2010.01385.x

Barlow, S. M., Boobis, A. R., Bridges, J., Cockburn, A., Dekant, W., Hepburn, P., Houben, G. F., König, J., Nauta, M. J., Schuermans, J., & Bánáti, D. (2015). The role of hazard- and risk-based approaches in ensuring food safety. In *Trends in Food Science and Technology* (Vol. 46, Issue 2, pp. 176–188). https://doi.org/10.1016/j.tifs.2015.10.007

Barnes, Amy Parkhurst, J. (2014). Can Global Health Policy be Depoliticized? A Critique of Global Calls for Evidence-Based Policy. In *The Handbook of Global Health Policy 84* (pp. 157–173). Wiley-Blackwell.

Bayley, C., & French, S. (2008). Designing a participatory process for stakeholder involvement in a societal decision. *Group Decision and Negotiation*, *17*(3), 195–210. https://doi.org/10.1007/s10726-007-9076-8

Beck, U. (1986). *Risikogesellschaft. Auf dem Weg in eine andere Moderne*. Suhrkamp. https://doi.org/10.1017/CBO9781107415324.004

Benford, D. (2000). *The Acceptable Daily Intake - A tool for ensuring food saftey*. ILSI Press. http://www.pac.gr/bcm/uploads/c2000acc_dai.pdf

BfR. (2016). *Spezial Pflanzenschutzmittel, BfR-Verbraucher Monitor 2016.*

Bieberstein, A., & Roosen, J. (2015). Gender differences in the meanings associated with food hazards: A means-end chain analysis. *Food Quality and Preference*, *42*, 165–176. https://doi.org/10.1016/j.foodqual.2014.11.003

Bitsch, V. (2005). Qualitative Research: A Grounded Theory Example and Evaluation Criteria. *Journal of Agribusiness*, *231*(Spring), 75–91.

Bogner, A., Littig, B., & Menz, W. (2009). *Interviewing Experts*. Palgrave Macmillan. https://doi.org/10.1057/9780230244276

Boholm, Å. (1998). Comparative studies of risk perception: A review of twenty years of research. *Journal of Risk Research*, *1*(2), 135–163. https://doi.org/10.1080/136698798377231

Boobis, A. R., Ossendorp, B. C., Banasiak, U., Hamey, P. Y., Sebestyen, I., & Moretto, A. (2008). Cumulative risk assessment of pesticide residues in food. *Toxicology Letters*, *180*(180), 137–150. https://doi.org/10.1016/j.toxlet.2008.06.004

Bouyer, M., Bagdassarian, S., Chaabanne, S., & Mullet, E. (2001). Personality correlates of risk perception. *Risk Analysis*, *21*(3), 457–465. https://doi.org/10.1111/0272-4332.213125

Bradbury, J. A. (1989). The policy implications of differing concepts of risk. *Science, Technology and Human Values*, *14*(4), 380–399. https://doi.org/doi.org/10.1177/016224398901400404

Brieger, W. (2017). The challenge of using and misusing insecticide-treated bed nets. *Africa Health* , *Sept.*, 13–15. https://doi.org/10.1186/1475-2875-13-464

Bronfman, N. C., Cifuentes, L. A., & Gutiérrez, V. V. (2008). Participant-focused analysis: explanatory power of the classic psychometric paradigm in risk perception. *Journal of Risk Research*, *11*(6), 735–753. https://doi.org/10.1080/13669870801967143

Bruhn, C., Diazknauf, K., Feldman, N., Harwood, J., Ho, G., Ivans, E., Kubin, L., Lamp, C., Marshall, M., Osaki, S., Stanford, G., Steinberg, Y., Valdez, I., Williamson, E., & Wunderlich, E. (1992). Consumer food safety concerns and interests in pesticide-related information. *Journal of Food Safety*, *12*(3), 253–262.

Bruhn, C. M., & Schutz, H. G. (1999). Consumer Food Safety Knowledge and Practices. *Journal of Food Safety*, *19*(1), 73–87. https://doi.org/10.1111/j.1745-4565.1999.tb00235.x

Buzby, J. C., Fox, J. A., Ready, R. C., & Crutchfield, S. R. (1998). Measuring Consumer Benefit of Food Safety Risk Reductions. *Joural of Agricultural and Applied Economics*, *30*(1), 69–82. https://doi.org/10.1017/S1074070800008063

BVL. (2017). *BVL - Wer macht was?* http://www.bvl.bund.de/DE/04_Pflanzenschutzmittel/01_Aufgaben/01_WerMachtWas/psm_WerMachtWas_node.html

Byrne, P. J., Gempesaw Ii, C. M., & Toensmeyer, U. C. (1991). An evaluation of consumer pesticide residue concerns and risk information sources. *Southern Journal of Agricultural Economics*, *December*. https://doi.org/10.1017/S0081305200018288

CAC/GL. (2007). *Working principles for risk analysis for food safety application by governments - CAC/GL 62-2007*.

Campbell, J. L., Quincy, C., Osserman, J., & Pedersen, O. K. (2013). Coding in-depth semistructured interviews: Problems of unitization and intercoder reliability and agreement. *Sociological Methods and Research*, *42*(3), 294–320. https://doi.org/10.1177/0049124113500475

Carson, R. (1962). *Silent spring*. https://doi.org/10.9774/gleaf.978-1-907643-44-6_4

Cartwright, N. (2006). Well Ordered Science: Evidence for Use. *Philosophy of Science*, *73*(5), 981–990. https://doi.org/10.1086/518803

Caswell, J. (1998). Valuing the benefits and costs of improved food safety and nutrition. *The Australian Journal of Agricultural and Resource Economics*, *42*(4), 409–424. https://doi.org/10.1111/1467-8489.00060

Cerroni, S., Notaro, S., & Shaw, W. D. (2013). How many bad apples are in a bunch? An experimental investigation of perceived pesticide residue risks. *Food Policy*, *41*, 112–123. https://doi.org/10.1016/j.foodpol.2013.04.012

Chatzopoulou, S., Eriksson, N. L., & Eriksson, D. (2020). Improving Risk Assessment in the European Food Saftey Authority: Lessons From The European Medicines Agency. *Frontiers in Plant Science*, *11*(349).

Chryssochoidis, G., Strada, A., & Krystallis, A. (2009). *Public trust in institutions and information sources regarding risk management and communication : towards integrating extant knowledge*. *12*(2), 137–185. https://doi.org/10.1080/13669870802637000

Clinton, C., & Sridhar, D. L. (2017). *Governing global health: who runs the world and why?* Oxfort University Press .

Cope, S., Frewer, L. J., Renn, O., & Dreyer, M. (2010). Potential methods and approaches to assess social impacts associated with food safety issues. *Food Control*, *21*, 1629–1637. https://doi.org/10.1016/j.foodcont.2009.05.008

Covello, V. T., & Mumpower, J. (1985). Risk Analysis and Risk Management: An Historical Perspective. *Risk Analysis*, *5*(2), 103–120. https://doi.org/10.1111/j.1539-6924.1985.tb00159.x

Crawford-Brown, D. J. (1999). *Risk-Based Environmental Decisions : Methods and Culture*. Springer US.

Cropper, M. L., Evans, W. N., Berardi, S. J., Ducla-Soares, M. M., & Portney, P. R. (1992). The determinants of pesticide regulation: A statistical analysis of EPA decision making. *Journal of Political Economy*, *100*(1), 175–197. https://doi.org/10.1086/261812

Cross, P., & Edwards-Jones, G. (2011). Variation in pesticide hazard from arable crop production in Great Britain from 1992 to 2008: An extended time-series analysis. *Crop Protection*, *30*(12), 1579–1585. https://doi.org/10.1016/j.cropro.2011.08.003

Cunha, L. M., Pinto De Moura, A., Lopes, Z., do Céu Santos, M., & Silva, I. (2010). Public perceptions of food-related hazards: an application to Portuguese consumers. *British Food Journal*, *112*(5), 522–543. https://doi.org/10.1108/eb011783

Cunningham, R. (2005). The ABC of GMOs, SPS and the WTO: an analysis of the application of the Agreement on Sanitary and Phytosanitary Measures within the context of biotechnology and international trade. *Southern Cross University Law Review*, *9*, 19–37. http://epubs.scu.edu.au/cgi/viewcontent.cgi?article=1100&context=law_pubs

D'Mello, J. P. F. (2003). *Food safety : contaminants and toxins*. CABI Pub. https://books.google.de/books/about/Food_Safety.html?id=VZ2TAIi4q54C&redir_esc=y

Damalas, C. A., & Eleftherohorinos, I. G. (2011). Pesticide exposure, safety issues, and risk assessment indicators.

International Journal of Environmental Research and Public Health, *8*(5), 1402–1419. https://doi.org/10.3390/ijerph8051402

Devos, Y., Elliott, K. C., Macdonald, P., McComas, K., Parrino, L., Vrbos, D., Robinson, T., Spiegelhalter, D., & Gallani, B. (2019). Conducting fit-for-purpose food safety risk assessments. *EFSA Journal*, *17*(S1), 1–16. https://doi.org/10.2903/j.efsa.2019.e170707

Disdier, L., Fontagne, L., & Mimouni, M. (2008). The impact of regulations on agricultural trade: evidence from the SPS and TBT agreements. *American Journal of Agricultural Economics*, *90*(2), 336–350. http://www.jstor.org/stable/30139588

Dobrow, M. J., Goel, V., & Upshur, R. E. G. (2004). Evidence-based health policy: Context and utilisation. *Social Science and Medicine*, *58*(1), 207–217. https://doi.org/10.1016/S0277-9536(03)00166-7

Dolan, C. B., BenYishay, A., Grépin, K. A., Tanner, J. C., Kimmel, A. D., Wheeler, D. C., & McCord, G. C. (2019). The impact of an insecticide treated bednet campaign on all-cause child mortality: A geospatial impact evaluation from the Democratic Republic of Congo. *PLOS ONE*, *14*(2). https://doi.org/10.1371/JOURNAL.PONE.0212890

Dosman, D. M., Adamowicz, W. L., & Hrudey, S. E. (2001). Socioeconomic Determinants of Health- and Food Safety-Related Risk Perceptions. *Risk Analysis*, *21*(2), 307–318. https://doi.org/10.1111/0272-4332.212113

Douglas, M., & Wildavsky, A. B. (1983). *Risk and culture : an essay on the selection of technical and environmental dangers*. University of California Press. https://www.ucpress.edu/book.php?isbn=9780520050631

Dreyer, M., & Renn, O. (2009). *Food Safety Governance: Integrating Science, Precaution and Public Involvement* (Vol. 15). Springer.

Dreyer, M., Renn, O., Cope, S., & Frewer, L. J. (2010). Including social impact assessment in food safety governance. *Food Control*, *21*(12), 1620–1628. https://doi.org/10.1016/j.foodcont.2009.05.007

Driesen, D. M. (2004). Distributing the Costs of Environmental, Health, and Safety Protection: The Feasability Principle, Cost-Benefit Analysis, and Regulatory Reform. *Boston College Environmental Affairs Law Review*, *32*(1). https://doi.org/10.2139/ssrn.478769

Drogué, S., & Demaria, F. (2012). Pesticide residues and trade, the apple of discord? *Food Policy*, *37*, 641–649. https://doi.org/10.1016/j.foodpol.2012.06.007

Dunlap, R. E., & Beus, C. E. (1992). Understanding public concerns about pesticides: an empirical examination. *Journal of Consumer Affairs*, *26*(2), 418–438. https://doi.org/10.1111/j.1745-6606.1992.tb00035.x

EC. (2017). *Initiative details - European Citizens' Initiative - European Commission*. http://ec.europa.eu/citizens-initiative/public/initiatives/successful/details/2017/000002/en?lg=en

EC. (2020). *Food Law General Principles | Food Safety*. https://ec.europa.eu/food/safety/general_food_law/principles_en

EC. (2021). *Glyphosate*. https://ec.europa.eu/food/plant/pesticides/glyphosate_en

EC, & EP. (2002). Regulation (EC) No 178/2002 of the European Parliament and the Council of 28 Januar 2002 laying down the general principles and requirements of food law, establishing the European Food Safety Authority and laying down procedures in matters of food safety. *Official Journal of the European Communities*, *1*(311). http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2002:031:0001:0024:en:PDF

EC, & EP. (2005a). Regulation (EC) No 2073/2005 of the Commission. In *Official Journal of the European Union* (Vol. 3651083, Issue 19). http://eur-lex.europa.eu/legal-content/DE/TXT/PDF/?uri=CELEX:02005R2073-20140601&from=EN

EC, & EP. (2005b). Regulation (EC) No 396/2005 of the European Parliament and of the Coucil. *Official Journal of the European Union*. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2005:070:0001:0016:en:PDF

EC, & EP. (2009). Regulation (EC) No 1107/2009 of the European Parliament and of the Coucil concerning the placing of plant protection products on the market and repealing Council Directives 79/117/EEC and 91/414/EEC THE. In *Official Journal of the European Union* (Vol. 309, Issue 1).

ECHA. (2017). *Classification - ECHA*. https://echa.europa.eu/regulations/clp/classification

EFSA. (2010). *Eurobarometer Spezial 354*. https://www.efsa.europa.eu/sites/default/files/corporate_publications/files/riskperceptionreport_de.pdf

EFSA. (2015a). *EFSA Strategy 2020*.

EFSA. (2015b). *Principles and process for dealing with data and evidence* (Vol. 13, Issue June). https://doi.org/10.2903/j.efsa.2015.4121

EFSA. (2016a). *EFSA Strategy 2020: Trusted science for safe food. Protecting consumers' health with independent scientific advice on the food chain*. https://doi.org/doi.org/10.2805/397609

EFSA. (2016b). *Pesticide residues in food: risk to consumers remains low*. https://www.efsa.europa.eu/en/press/news/161026

EFSA. (2017). *Pesticide evaluation: Tools | European Food Safety Authority*. https://www.efsa.europa.eu/de/applications/pesticides/tools

EFSA. (2018). *Risk communication: tools and approaches* (Issue October). https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwjM_bPs8OLqAhWKxqQKHcsrBg0QFjABegQIBBAB&url=https%3A%2F%2Fwww.efsa.europa.eu%2Fsites%2Fdefault%2Ffiles%2Fevent%2Fmb181010%2Fmb181010-i7-p.pdf&usg=AOvVaw20hIt0sbD-aa5VNI5sFnrP

EFSA. (2019). *2019 Eurobarometer on Food Safety in the EU*. https://www.efsa.europa.eu/en/corporate/pub/eurobarometer19

EFSA. (2020). *Organisational Structure EFSA*. EFSA Webpage. https://www.efsa.europa.eu/sites/default/files/assets/orgchart.pdf

EFSA, & ECDC. (2018). The European Union summary report on trends and sources of zoonoses, zoonotic agents and food-borne outbreaks in 2017. *EFSA Journal*, *16*(12). https://doi.org/10.2903/j.efsa.2018.5500

EFSA, Medina-Pastor, P., & Triacchini, G. (2020). The 2018 European Union report on pesticide residues in food. In *EFSA Journal* (Vol. 18, Issue 4). https://doi.org/10.2903/j.efsa.2020.6057

EFSA Panel on BIOHAZ. (2012). Scientific opinion on the development of a risk ranking framework on biological hazards. *EFSA Journal*, *10*(6). https://doi.org/10.2903/j.efsa.2012.2724.

Eom, Y. S. (1994). Pesticide Residue Risk and Food Safety Valuation: A Random Utility Approach. *American Journal of Agricultural Economics*, *76*(4), 760–771. https://doi.org/10.2307/1243737

Erdem, S., Rigby, D., & Wossink, A. (2012). Using best-worst scaling to explore perceptions of relative responsibility for ensuring food safety. *Food Policy*, *37*(6), 661–670. https://doi.org/10.1016/j.foodpol.2012.07.010

Erlacher, E., & Wang, M. (2011). Regulation (EC) No. 1107/2009 and upcoming challenges for exposure assessment of plant protection products - Harmonisation or national modelling approaches? *Environmental Pollution*, *159*(12), 3357–3363. https://doi.org/10.1016/j.envpol.2011.08.036

Essaji, A. (2008). Technical regulations and specialization in international trade. *Journal of International Economics*, *76*(2), 166–176. https://doi.org/10.1016/j.jinteco.2008.06.008

Evidence-Based Policymaking Collaborative. (2019). *Principles of Evidence-Based Policymaking | Evidence-Based Policymaking Collaborative*. https://www.evidencecollaborative.org/principles-evidence-based-policymaking

Fantke, P., Friedrich, R., & Jolliet, O. (2012). Health impact and damage cost assessment of pesticides in Europe. *Environment International*, *49*, 9–17. https://doi.org/10.1016/j.envint.2012.08.001

FAO/WHO. (2020). *About Codex*. http://www.fao.org/fao-who-codexalimentarius/about-codex/en/#c453333

FAO, & WHO. (1997). *Risk management and food safety* (Issue 65). http://www.fao.org/3/W4982E/w4982e00.htm

FAO, & WHO. (2009). *Guidelines for risk characterization of microbiological hazards in food*. http://www.fao.org/3/i1134e/i1134e00.htm

Fazil, A. M., Rajic, A., Sanchez, J., & McEwen, S. A. (2008). Choices, choices: the application of multi-criteria decision analysis to a food safety decision-making problem. In *Journal of Food Protection* (Vol. 71, Issue 11, pp. 2323–2333). https://doi.org/10.4315/0362-028X-71.11.2323

Fife-Schaw, C., & Rowe, G. (1996). Public perceptions of everyday food hazards: a psychometric study. *Risk Analysis*, *16*(4), 487–500. https://doi.org/10.1111/j.1539-6924.1996.tb01095.x

Finn, A., & Louviere, J. J. (1992). Determing the Appropriate Response to Evidence of Public Concern: The Case of Food Safety. *Journal of Public Policy & Marketing*, *11*(2), 12–25. https://doi.org/10.2307/30000270

Finucane, M., Alhakami, A., Slovic, P., & Johnson, S. M. (2000). The affect heuristic in judgements of risks and benefits. *Journal of Behavioral Decision Making*, *13*, 1–17. http://www-abc.mpib-berlin.mpg.de/users/r20/finucane00_the_affect_heuristic.pdf

Fischhoff, B. (1994). Acceptable risk: A conceptual proposal. *RISK: Health, Safety & Environment*, *5*(1), 3.

Fischhoff, B., & Hope, C. (1984). Defining risk. *Policy Sciences*, *17*, 123–139. https://doi.org/https://doi.org/10.1007/BF00146924

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

Flick, U. (2011). *Triangulation*. Springer. https://www.springer.com/de/book/9783531181257

Florax, R. J. G. M., Travisi, C. M., & Nijkamp, P. (2005). A meta-analysis of the willingness to pay for reductions in pesticide risk exposure. *European Review of Agricultural Economics*, *32*(4 SPEC. ISS.), 441–467. https://doi.org/10.1093/erae/jbi025

Focker, M., & van der Fels-Klerx, H. J. (2020). Economics applied to food safety. *Current Opinion in Food Science*, *36*, 18–23. https://doi.org/10.1016/j.cofs.2020.10.018

Fraiberg, J. D., & Trebilcock, M. J. (1998). Risk Regulation: Technocratic and Democratic Tools for Regulatory Reform. *Mc Gill Law Journal*, *43*, 835–888. http://www.lawjournal.mcgill.ca/userfiles/other/2547410-43.FraibergTrebilcock.pdf

Fraser, N., Druce, N., Delgado, J., WHO, UNICEF/UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases., & et al. (2006). *Partnerships for malaria control : engaging the formal and informal private sectors : a review / commissioned by the UNICEF/UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases (TDR), in collaboration with the Working Group on Financing and Resources of the Roll Back Malaria Partnership, chaired by The World Bank*. https://apps.who.int/iris/handle/10665/43577

Frewer, L. J., Fischer, A. R. H. H., Brennan, M., Bánáti, D., Lion, R., Meertens, R. M., Rowe, G., Siegrist, M., Verbeke, W., & Vereijken, C. M. J. L. J. L. (2016). Risk/Benefit Communication about Food—A Systematic Review of the Literature. *Critical Reviews in Food Science and Nutrition*, *56*(10), 1728–1745. https://doi.org/10.1080/10408398.2013.801337

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

Galison, P. (2010). Trading with the enemy. In Michael E. Groman (Ed.), *Trading Zones and Interactional Expertise : Creating New Kinds of Collaboration* (Vol. 21, Issue 1, pp. 147–175). The MIT Press. https://doi.org/10.1017/S0018246X00012152

Gaskell, G., Hohl, K., & Gerber, M. M. (2017). Do closed survey questions overestimate public perceptions of food risks? *Journal of Risk Research*, *20*(8), 1038–1052. https://doi.org/10.1080/13669877.2016.1147492

Gioia, D. A., Corley, K. G., & Hamilton, A. L. (2012). Seeking qualitative rigor in inductive research. *Organizational Research Methods*, *16*(1), 15–31. https://doi.org/10.1177/1094428112452151

Gläser, J., & Laudel, G. (2010). *Expert interviews and qualitative content analysis*. VS Verlag für Sozialwissenschaften.

Goodman, CA, Coleman, P., & Mills, A. (1999). Cost-effectiveness of malaria control in sub-Saharan Africa. *Lancet*, *31*(354 (9176)), 378–385. https://doi.org/10.1016/s0140-6736(99)02141-8

Graham, J. D., & Wiener, J. B. (1995). *Risk versus risk: Tradeoffs in protecting health and the environment*. Harvard University Press. https://doi.org/Doi 10.2307/2151986

Gray, G. M., & Hammitt, J. K. (2000). Risk/risk trade-offs in pesticide regulation: An exploratory analysis of the public health effects of a ban on organophosphate and carbamate pesticides. *Risk Analysis*, *20*(5), 665–680. https://doi.org/10.1111/0272-4332.205060

Gruev-Vintila, A., & Rouquette, M. L. (2007). Social thinking about collective risk: How do risk-related practice and personal involvement impact its social representations? *Journal of Risk Research*, *10*(4), 555–581. https://doi.org/10.1080/13669870701338064

Grunert, K. G. (2005). Food quality and safety: consumer perception and demand. *European Review of Agricultural Economics*, *32*(3), 369–391. https://doi.org/10.1093/eurrag/jbi011

Gruszczynski, L. (2007). Science in the Process of Risk Regulation under the WTO Agreement on Sanitary and Phytosanitary Measures. *German Law Journal*, *07*(4). https://static1.squarespace.com/static/56330ad3e4b0733dcc0c8495/t/56b87452c2ea51d6b7d1e472/1454928979009/GLJ_Vol_07_No_04_Gruszczynski.pdf

Guimarães, R. M., Asmus, C. I. R. F., & Meyer, A. (2007). DDT reintroduction for malaria control: the cost-benefit debate for public health. *Cadernos de Saude Publica*, *23*(12), 2835–2844. https://doi.org/10.1590/S0102-311X2007001200004

Guyton, K. Z., Loomis, D., Grosse, Y., El Ghissassi, F., Benbrahim-Tallaa, L., Guha, N., Scoccianti, C., Mattock, H., Straif, K., Blair, A., Fritschi, L., McLaughlin, J., Sergi, C. M., Calaf, G. M., Le Curieux, F., Baldi, I., Forastiere, F., Kromhout, H., 't Mannetje, A., … Zeise, L. (2015). Carcinogenicity of tetrachlorvinphos, parathion, malathion, diazinon, and glyphosate. *The Lancet Oncology*, *16*(5), 490–491. https://doi.org/10.1016/S1470-2045(15)70134-8

Haddaway, N. R., & Bilotta, G. S. (2016). Systematic reviews: Separating fact from fiction. *Environment International*, *92–93*, 578–584. https://doi.org/10.1016/j.envint.2015.07.011

Haidt, J. (2012). *The righteous mind : why good people are divided by politics and religion*. Pantheon Books.

Handford, C. E., Elliott, C. T., & Campbell, K. (2015). A review of the global pesticide legislation and the scale of challenge in reaching the global harmonization of food safety standards. *Integrated Environmental Assessment and Management*, *11*(4), 525–536. https://doi.org/10.1002/ieam.1635

Hansen, J., Holm, L., Frewer, L., Robinson, P., & Sandoe, P. (2003). Beyond the knowledge deficit: Recent research into lay and expert attitudes to food risks. In *Appetite* (Vol. 41, Issue 2, pp. 111–121). https://doi.org/10.1016/S0195-

6663(03)00079-5

Hansen, S. F., Von Krauss, M. K., & Tickner, J. A. (2008). The precautionary principle and risk-risk tradeoffs. *Journal of Risk Research*, *11*(4), 423–464. https://doi.org/10.1080/13669870801967192

Harper, C. R., & Zilberman, D. (1992). Pesticides and Worker Safety. *American Journal of Agricultural Economics*, *74*(1), 68–78. https://doi.org/10.2307/1242991

Hassauer, C., & Roosen, J. (2019). Which criteria do consumers use to evaluate safe food? *AAEA Annual Meeting*.

Hassauer, C., & Roosen, J. (2020). Toward a conceptual framework for food safety criteria: Analyzing evidence practices using the case of plant protection products. *Safety Science*, *127*, 104683. https://doi.org/10.1016/j.ssci.2020.104683

Hayek, F. A. (1945). The Use of Knowledge in Society. *The American Economic Review*, *35*(4), 519–530. https://doi.org/10.2307/1809376

Henning, K. J., Freyberg, S., Stehfest, S., Kaus, S., & Böl, G.-F. (2014). *EU-Almanach Lebensmittelsicherheit*. http://www.efsa.europa.eu/en/scdocs/doc/art36listg.pdf

Henson, S., & Loader, R. (2001). Barriers to agricultural exports from developing countries: The role of sanitary and phytosanitary requirements. *World Development*, *29*(1), 85–102. https://doi.org/10.1016/S0305-750X(00)00085-1

Henson, Spencer, & Traill, B. (1993). The demand for food safety. Market imperfections and the role of government. *Food Policy*, *18*(2), 152–162. https://doi.org/10.1016/0306-9192(93)90023-5

Herges, L., Kaus, S., Böl, G.-F., & Gollnick, N. (2017). *EU-Almanach Lebensmittelsicherheit*. http://www.bfr.bund.de/de/publikation/eu_almanach-192686.html

Hessing, A., Schneider, R. G., Gutierrez, A., Silverberg, R., & Michael, S. (2015). *The Cost of Food Safety*. https://edis.ifas.ufl.edu/fs270

Hicks, J. R. (1939). The Foundations of Welfare Economics. *The Economic Journal*, *49*(196), 696. https://doi.org/10.2307/2225023

Hill, J., Lines, J., & Rowland, M. (2006). Insecticide-treated nets. *Advances in Parasitology*, *61*, 77–128. https://doi.org/10.1016/S0065-308X(05)61003-2

Hirschman, A. (1971). *Exit, voice and loyalty: Responses to decline in firms, organisation, and states*. Harvard University Press.

Hobbs, J. E., Bailey, D., Dickinson, D. L., & Haghiri, M. (2005). Traceability in the Canadian Red Meat Sector: Do Consumers Care? *Canadian Journal of Agricultural Economics*, *53*(1), 47–65. https://doi.org/10.1111/j.1744-7976

Hoffmann, S. (2010). Ensuring food safety around the globe: The many roles of risk analysis from risk ranking to microbial risk assessment: Introduction to special series. *Risk Analysis*, *30*(5), 711–714. https://doi.org/10.1111/j.1539-6924.2010.01437.x

Homan, J., Petts, J., Pollard, S., & Twigger-Ross, C. (2001). Participatory risk assessment for environmental decision-making. *Values in Decisions on Risk. 2. VALDOR Symposium*. https://www.osti.gov/etdeweb/biblio/20412454

Hornibrook, S. A., McCarthy, M., & Fearne, A. (2005). Consumers' perception of risk: The case of beef purchases in Irish supermarkets. *International Journal of Retail & Distribution Management*, *33*(10), 701–715. https://doi.org/10.1108/09590550510622263

Hussain, M., & Dawson, C. (2013). Economic impact of food safety outbreaks on food businesses. *Foods*, *2*(4), 585–589. https://doi.org/10.3390/foods2040585

Ingram, J. (2011). A food systems approach to researching food security and its interactions with global environmental change. *Food Security*, *3*(4), 417–431. https://doi.org/10.1007/s12571-011-0149-9

Initiative Volksbegehren Artenvielfalt. (2019). *Volksbegehren Artenvielfalt*. Web Page. https://volksbegehren-artenvielfalt.de

IoM, & NRC. (2003). *Scientific criteria to ensure safe food*. National Academies Press. https://doi.org/10.17226/10690

Jabareen, Y. (2009). Building a Conceptual Framework: Philosophy, Definitions, and Procedure. *International Journal of Qualitative Methods*, *8*(4). https://journals.library.ualberta.ca/ijqm/index.php/IJQM/article/viewFile/6118/5892

James L. A. Webb, J. (2014). The Long Struggle against Malaria in Tropical Africa. In *The Long Struggle against Malaria in Tropical Africa*. Cambridge University Press. https://doi.org/10.1017/CBO9781107280625

Jasanoff, S. (1990). *The fifth branch : science advisers as policymakers*. Harvard University Press.

Johansen, I. L., & Rausand, M. (2015). Ambiguity in risk assessment. *Safety Science*, *80*, 243–251. https://doi.org/10.1016/j.ssci.2015.07.028

Kabasenche, W. P., & Skinner, M. K. (2014). DDT, epigenetic harm, and transgenerational environmental justice. *Environmental Health: A Global Access Science Source*, *13*(62). https://doi.org/10.1186/1476-069X-13-

62/TABLES/1

Kahan, D. M. (2016). The politically motivated reasoning paradigm. *Emerging Trends in Social & Behavioral Sciences*, 1–15. https://doi.org/10.1002/9781118900772

Kahan, D. M., Braman, D., Gastil, J., Slovic, P., & Mertz, C. K. (2007). Culture and identity-protective cognition: Explaining the white male effect in risk perception. *Journal of Empirical Law Studies*, *4*(3), 465–505. https://doi.org/10.1111/j.1740-1461.2007.00097.x

Kaldor, N. (1939). Welfare Propositions of Economics and Interpersonal Comparisons of Utility. *The Economic Journal*, *49*(195), 549. https://doi.org/10.2307/2224835

Kaplan, L. B., Szybillo, G. J., & Jacoby, J. (1974). Components of perceived risk in product purchase: A cross-validation. *Journal of Applied Psychology*, *59*(3), 287–291. https://doi.org/10.1037/h0036657

Kaptan, G., Fischer, A. R. H., & Frewer, L. J. (2017). Extrapolating understanding of food risk perceptions to emerging food safety cases. *Journal of Risk Research*, *9877*(February), 1–23. https://doi.org/10.1080/13669877.2017.1281330

Karabelas, A. J., Plakas, K. V, Solomou, E. S., Drossou, V., & Sarigiannis, D. A. (2009). Impact of European legislation on marketed pesticides - A view from the standpoint of health impact assessment studies. In *Environment International* (Vol. 35, Issue 7, pp. 1096–1107). https://doi.org/10.1016/j.envint.2009.06.011

Kasperson, R., Renn, O., Slovic, P., Brown, H., Emel, J., Goble, R., Kasperson, J., & Ratick, S. (1988). The Social Amplification of Risk-A Conceptual Framework. *Risk Analysis*, *8*(2), 177–187. https://doi.org/10.1111/j.1539-6924.1988.tb01168.x

Kawasaki, K., & Lichtenberg, E. (2015). Quality versus Quantity Effects of Pesticides: Joint Estimation of Quality Grade and Crop Yield. *AAEA Annual Meeting*. https://ageconsearch.umn.edu/bitstream/204848/2/Quality versus Quantity Effects of Pesticides Joint Estimation of Quality Grade and Crop Yield AAEA 150515 AgEcon Search.pdf

Kelly, T. (2016). Evidence. In *Stanford Encyclopedia of Philosophy*.

Kennedy, M. C., Van Der Voet, H., Roelofs, V. J., Roelofs, W., Glass, C. R., De Boer, W. J., Kruisselbrink, J. W., & Hart, A. D. M. (2015). New approaches to uncertainty analysis for use in aggregate and cumulative risk assessment of pesticides. *Food and Chemical Toxicology*, *79*, 54–64. https://doi.org/10.1016/j.fct.2015.02.008

Kerr, W. A. (2009). Political Precaution, Pandemics and Protectionism and Trade Policy. *The Estey Centre Journal of International Law and Trade Policy*, *10*(2), 1–14. http://ageconsearch.umn.edu/record/52231/files/kerr10-2.pdf

Kher, S. V., De Jonge, J., Wentholt, M. T. A., Deliza, R., de Andrade, J. C., Cnossen, H. J., Luijckx, N. B. L., & Frewer, L. J. (2013). Consumer perceptions of risks of chemical and microbiological contaminants associated with food chains: A cross-national study. *International Journal of Consumer Studies*, *37*(1), 73–83. https://doi.org/10.1111/j.1470-6431.2011.01054.x

Kinchy, A. J. (2010). Anti-genetic engineering activism and scientized politics in the case of "contaminated" Mexican maize. *Agriculture and Human Values*, *27*(4), 505–517. https://doi.org/10.1007/s10460-009-9253-2

Klinke, A., & Renn, O. (2002). A new approach to risk evaluation and management: Risk-based, precaution-based, and discourse-based strategies. *Risk Analysis*, *22*(6), 1071–1094. https://doi.org/10.1111/1539-6924.00274

Knight, A., & Warland, R. (2004). The relationship between sociodemographics and concern about food safety issues. *Journal of Consumer Affairs*, *28*(1), 107–120. https://doi.org/10.1111/j.1745-6606.2004.tb00467.x

König, A. (2007). Towards safer foods and more democratic decisions. Is this a contradictory goal? *OCL - Oleagineux Corps Gras Lipides*, *14*(2), 92–99. https://doi.org/10.1684/ocl.2007.0130

König, A., Kuiper, H. A., Marvin, H. J. P., Boon, P. E., Busk, L., Cnudde, F., Cope, S., Davies, H. V, Dreyer, M., Frewer, L. J., Kaiser, M., Kleter, G. A., Knudsen, I., Pascal, G., Prandini, A., Renn, O., Smith, M. R., Traill, B. W., Voet, H. van der, … Wentholt, M. T. A. (2010). The SAFE FOODS framework for improved risk analysis of foods. *Food Control*, *21*(12), 1566–1587. https://doi.org/10.1016/j.foodcont.2010.02.012

Kouyaté, B., Sie, A., Yé, M., De Allegri, M., & Müller, O. (2007). The great failure of malaria control in Africa: a district perspective from Burkina Faso. *PLoS Medicine*, *4*(6). https://doi.org/10.1371/JOURNAL.PMED.0040127

Kramer Professor, R., Dickinson, K. L., Anderson Assistant Professor, R. M., Fowler Associate Professor, V. G., Lynn Miranda Associate Professor, M., Mutero Head, C. M., Saterson, K. A., & Wiener Perkins Professor, J. B. (2009). Using Decision Analysis to Improve Malaria Control Policy Making. *Health Policy*, *92*(0), 133–140. https://doi.org/10.1016/j.healthpol.2009.02.011

Kraus, N., Malmfors, T., & Slovic, P. (1992). Intuitive Toxicology : Expert and Lay Judgments of Chemical Risks. *Risk Analysis*, *12*(2), 215–232.

Krezanoski, P. J. (2016). Delivering insecticide-treated nets for malaria prevention: innovative strategies. *Research and Reports in Tropical Medicine*, *7*, 39–47. https://doi.org/10.2147/RRTM.S83173

Lancaster, K., Seear, K., Treloar, C., & Ritter, A. (2017). The productive techniques and constitutive effects of 'evidence-based policy' and 'consumer participation' discourses in health policy processes. *Social Science and Medicine*, *176*,

60–68. https://doi.org/10.1016/j.socscimed.2017.01.031

Leiss, W., & Powell, D. (2004). *Mad cows and mothers milk: The perils of poor risk communication*. McGill-Queen's University Press. https://www.scopus.com/inward/record.uri?eid=2-s2.0-84897308799&partnerID=40&md5=11a8b45c350529bb0a158952d20623e7

Lichtenberg, E., Parker, D. D., & Zilberman, D. (1988). Marginal Analysis of Welfare Costs of Environmental Policites: The Case of Pesticide Regulation. *American Journal of Agricultural Economics*, *70*(4), 867–874. https://doi.org/10.2307/1241928

Lin, C.-T. J. (1995). Demographic and Socioeconomic Influences on the Importance of Food Safety in Food Shopping. *Agricultural and Resource Economics Review*, *24*, 191–198. https://doi.org/10.1017/S1068280500008832

Lineback, D. R., Pirlet, A., Van Der Kamp, J. W., & Wood, R. (2009). Globalization, food safety issues & role of international standards. *Quality Assurance and Safety of Crops and Foods*, *1*(1), 23–27. https://doi.org/10.1111/j.1757-837X.2009.00005.x

Lobb, A. E., Mazzocchi, M., & Traill, W. B. (2007). Modelling risk perception and trust in food safety information within the theory of planned behaviour. *Food Quality and Preference*, *18*(2), 384–395. https://doi.org/10.1016/j.foodqual.2006.04.004

Loewenstein, G. F., Weber, E. U., Hsee, C. K., & Welch, N. (2001). Risk as feelings. *Psychological Bulletin*, *127*(2), 267–286. https://doi.org/10.1037/0033-2909.127.2.267

Lofstedt, R. E. (2006). How can we Make Food Risk Communication Better: Where are we and Where are we Going? *Journal of Risk Research*, *9*(8), 869–890. https://doi.org/10.1080/13669870601065585

Lusk, J. L. (2013). *The Food Police*. Crown Forum.

Lusk, J. L., Roosen, J., & Bieberstein, A. (2014). Consumer Acceptance of New Food Technologies: Causes and Roots of Controversies. *Annual Review of Resource Economics*, *6*, 381–405. https://doi.org/10.1146/annurev-resource-100913-012735

Lynch, M. P. (2012). *In Praise of Reason: Why Rationality Matters for Democracy*. MIT Press.

Maclachlan, D. J., & Hamilton, D. (2010). Estimation methods for Maximum Residue Limits for pesticides. *Regulatory Toxicology and Pharmacology*, *58*, 208–218. https://doi.org/10.1016/j.yrtph.2010.05.012

Mahon, D., & Cowan, C. (2004). Irish consumers' perception of food safety risk in minced beef. *British Food Journal*, *106*(4), 301–312. https://doi.org/10.1108/00070700410529564

Manfreda, G., & De Cesare, A. (2014). The challenge of defining risk-based metrics to improve food safety: Inputs from the BASELINE project. *International Journal of Food Microbiology*, *184*, 2–7. https://doi.org/10.1016/j.ijfoodmicro.2014.01.013

Maria Travisi, C., Nijkamp, P., & Vindigni, G. (2006). Pesticide risk valuation in empirical economics: A comparative approach. *Ecological Economics*, *56*(4), 455–474. https://doi.org/10.1016/j.ecolecon.2004.06.026

Marris, C., Langford, I., Saunderson, T., & O'Riordan, T. (1997). Exploring the "psychometric paradigm": Comparisons between aggregate and individual analyses. *Risk Analysis*, *17*(3), 303–312. https://doi.org/10.1111/j.1539-6924.1997.tb00868.x

Mccluskey, J. J., & Swinnen, J. F. M. (2004). Political Economy of the Media and Consumer Perceptions of Biotechnology. *American Journal of Agricultural Economics*, *86*(5), 1230–1237. http://content.ebscohost.com/ContentServer.asp?T=P&P=AN&K=15126072&S=R&D=bth&EbscoContent=dGJyMMTo50Seprc4yOvsOLCmr0%2BeprFSrq%2B4Sa6WxWXS&ContentCustomer=dGJyMPGvsE2yqLNLuePfgeyx44Dt6fIA

Membré, J. M., & Guillou, S. (2016). Lastest developments in foodborne pathogen risk assessment. *Current Opinion in Food Science*, *8*, 120–126. https://doi.org/10.1016/j.cofs.2016.04.011

Mempel, H. (2015). *Wesentliche Ergebnisse aus der ALDI SÜD Rückstandsdatenbank*. https://unternehmen.aldi-sued.de/fileadmin/fm-dam/company_photos/US_Verantwortung/Downloads/Studie_Pflanzenschutzmittel_Kurzfassung_Prof._Dr._Heike_Mempel.pdf

Merriam-Webster. (2019). *Definition of Evaluate*. https://www.merriam-webster.com/dictionary/evaluate

Merriam-Webster. (2021). *Definition of relevant*. Dictionary. https://www.merriam-webster.com/dictionary/relevant

Meuser, M., & Nagel, U. (1991). ExpertInneninterviews - vielfach verprobt, wenig bedacht. Ein Beitrag zur qualitativen Methodendiskussion. In *Qualitativ-empirische Sozialforschung*. https://doi.org/10.1007/978-3-322-97024-4

Meuser, M., & Nagel, U. (2009). The expert interview and changes in knowledge production. In *Interviewing Experts* (pp. 17–42). Palgrave Macmillan.

Miles, S., Brennan, M., Kuznesof, S., Ness, M., Ritson, C., & Frewer, L. J. (2004). Public worry about specific food safety

issues. *British Food Journal*, *106*(1), 9–22. https://doi.org/10.1108/00070700410515172

Miles, S., & Frewer, L. J. (2001). Investigating specific concerns about different food hazards. *Food Quality and Preference*, *12*(1), 47–61. https://doi.org/10.1016/S0950-3293(00)00029-X

Millstone, E. (2009). Science, risk and governance: Radical rhetorics and the realities of reform in food safety governance. *Research Policy*, *38*(4), 624–636. https://doi.org/10.1016/j.respol.2009.01.012

Moore-Sheeley, K. (2017). *Nothing but Nets: The History of Insecticide-Treated Nets in Africa* [John Hopkins University]. https://jscholarship.library.jhu.edu/handle/1774.2/60905

Morvillo, M. (2020). Glyphosate effect: Has the Glyphosate controversy affected the EU's regulatory epistemology? *European Journal of Risk Regulation*, 1–14. https://doi.org/10.1017/err.2020.11

Mostafalou, S., & Abdollahi, M. (2013). Pesticides and human chronic diseases: Evidences, mechanisms, and perspectives. *Toxicology and Applied Pharmacology*, *268*(2), 157–177. https://doi.org/10.1016/j.taap.2013.01.025

Mullet, E., Duquesnoy, C., Raiff, P., Fahrasmane, R., & Namur, E. (1993). The Evaluative Factor of Risk Perception. *Journal of Applied Social Psychology*, *23*(19), 1594–1605. https://doi.org/10.1111/j.1559-1816.1993.tb01050.x

Nabarro, D. N., & Tayler, E. M. (1998). The "Roll Back Malaria" Campaign. *Science*, *280*(5372), 2067–2068. https://doi.org/10.1126/SCIENCE.280.5372.2067

Nau, H., Steinberg, P., & Kietzmann, M. (2002). *Lebensmitteltoxikologie: Rückstände und Kontaminanten: Risiken und Verbraucherschutz*. Parey Buchverlag.

Nayga, R. M. (1996). Sociodemographic Influences on Consumer Concern for Food Safety: The Case of Irradiation, Antibiotics, Hormones, and Pesticides. *Applied Economic Perspectives and Policy*, *18*(3), 467–475. https://doi.org/10.2307/1349629

Newsome, R., Tran, N., Paoli, G. M., Jaykus, L. A., Tompkin, B., Miliotis, M., Ruthman, T., Hartnett, E., Busta, F. F., Petersen, B., Shank, F., McEntire, J., Hotchkiss, J., Wagner, M., & Schaffner, D. W. (2009). Development of a risk-ranking framework to evaluate potential high-threat microorganisms, toxins, and chemicals in food. *Journal of Food Science*, *74*(2). https://doi.org/10.1111/j.1750-3841.2008.01042.x

Nordlander, K., Simon, C., & Pearson, H. (2010). Hazard v . Risk in EU Chemicals Regulation. *European Journal of Risk Regulation*, *3*, 239–251.

NRC. (1983). *Risk assessment in the federal government*. National Academies Press. https://doi.org/10.17226/366

NRC. (2009). Science and decisions: Advancing risk assessment. In *Science and Decisions: Advancing Risk Assessment*. National Academies Press. https://doi.org/10.17226/12209

Nygard, B., & Storstad, O. (1998). De-globalization of food markets? Consumer perceptions of safe food: the case of Norway. *Sociologia Ruralis*, *38*(1), 35–53. https://doi.org/10.1111/1467-9523.00062

OECD. (1993). *Safety evaluation of foods derived by modern biotechnology*. http://dbtbiosafety.nic.in/guideline/OACD/Concepts_and_Principles_1993.pdf

Oxford University Press. (2017). *criterion | Definition of criterion in US English by Oxford Dictionaries*. https://en.oxforddictionaries.com/definition/us/criterion

Parks, R. (2014). The Rise, Critique and Persistence of the DALY in Global Health. *The Columbia University Journal of Global Health*, *4*(1), 28–32. https://doi.org/10.7916/THEJGH.V4I1.4893

Patton, M. Q. (2015). *Qualitative research & evaluation methods*. SAGE. https://doi.org/10.1016/j.foreco.2006.01.032

Pease, C. K., & Gentry, R. P. (2016). Systematic review in chemical risk assessment — A chemical industry perspective. *Environment International*, *92–93*, 574–577. https://doi.org/10.1016/j.envint.2015.12.007

Pennings, J. M. E., Wansink, B., & Meulenberg, M. T. G. (2002). A note on modeling consumer reactions to a crisis: The case of the mad cow disease. *International Journal of Research in Marketing*, *19*(1), 91–100. https://doi.org/10.1016/S0167-8116(02)00050-2

Peters, E. M., Burraston, B., & Mertz, C. K. (2004). An emotion-based model of risk perception and stigma susceptibility: Cognitive appraisals of emotion, affective reactivity, worldviews, and risk perceptions in the generation of technological stigma. In *Risk Analysis* (Vol. 24, Issue 5, pp. 1349–1367). https://doi.org/10.1111/j.0272-4332.2004.00531.x

Peterson, M. (2006). The precautionary principle is incoherent. *Risk Analysis*, *26*(3), 595–601. https://doi.org/10.1111/j.1539-6924.2006.00781.x

Peterson, M. (2017). Yes, The Precautionary Principle Is Incoherent. In *Risk Analysis* (Vol. 37, Issue 11, pp. 2035–2038). https://doi.org/10.1111/risa.12783

Peterson, R. K. D., Barber, L. M., & Schleier, J. J. (2011). Net Risk: A Risk Assessment of Long-Lasting Insecticide Bed Nets Used for Malaria Management. *The American Journal of Tropical Medicine and Hygiene*, *84*(6), 956. https://doi.org/10.4269/AJTMH.2011.11-0016

Pfeiffer, J., & Nichter, M. (2008). What Can Critical Medical Anthropology Contribute to Global Health? *Medical Anthropology Quarterly*, *22*(4), 410–415. https://doi.org/10.1111/J.1548-1387.2008.00041.X

Pollak, R. A. (1998). *Imagined Risks and Cost-Benefit Analysis*. *88*(2), 376–380. http://www.jstor.org/stable/116951

Pollan, M. (2006). *The Omnivore's Dilemma*. The Penguin Press.

Porter, T. M. (1997). *Trust in numbers : the pursuit of objectivity in science and public life*. Princeton University Press.

Portier, C. J., Armstrong, B. K., Baguley, B. C., Baur, X., Belyaev, I., Bellé, R., Belpoggi, F., Biggeri, A., Bosland, M. C., Bruzzi, P., Budnik, L. T., Bugge, M. D., Burns, K., Calaf, G. M., Carpenter, D. O., Carpenter, H. M., López-carrillo, L., Clapp, R., Cocco, P., … Zhou, S.-F. (2016). Differences in the carcinogenic evaluation of glyphosate between the International Agency for Research on Cancer (IARC) and the European Food Safety Authority (EFSA). *J Epidemiol Community Health*, 1–5. https://doi.org/10.1136/jech-2015-207005

Purdy, M., Robinson, M., Wei, K., & Rublin, D. (2013). The economic case for combating malaria. *The American Journal of Tropical Medicine and Hygiene*, *89*(5), 819–823. https://doi.org/10.4269/AJTMH.12-0689

Raude, J., Fischler, C., Setbon, M., & Flahault, A. (2005). Scientist and public responses to BSE-related risk: A comparative study. In *Journal of Risk Research* (Vol. 8, Issues 7–8, pp. 663–678). https://doi.org/10.1080/13669870500194825

RBM Partnership to End Malaria. (2015a). *Action and Investment to defeat Malaria 2016–2030*. https://endmalaria.org/sites/default/files/RBM_AIM_Report_0.pdf

RBM Partnership to End Malaria. (2015b). *Evidence for Advocacy. Key Statitics on the Fight for Malaria*. http://www.makingmalariahistory.org/wp-content/uploads/2015/06/Malaria_Evidence-for-Advocacy_April_2015.pdf

RBM Partnership to End Malaria. (2015c). *Multisectoral Action Framework for Malaria*. https://endmalaria.org/sites/default/files/9_Multisectoral-Action-Framework-for-Malaria.pdf

RBM Partnership to End Malaria. (2017). *Annual Report*. https://endmalaria.org/sites/default/files/RBM Annual Report_web.pdf

RBM Partnership to End Malaria. (2018a). *Annual Report*. https://endmalaria.org/sites/default/files/RBM Annual Report 2018_EN.pdf

RBM Partnership to End Malaria. (2018b). *RBM Partnership Strategic Plan 2018–2020*. https://endmalaria.org/sites/default/files/RBM-Strategic-Plan-digital-JW-220218.pdf

Redmond, E. C., & Griffith, C. J. (2004). Consumer perceptions of food safety risk, control and responsibility. *Appetite*, *43*(3), 309–313. https://doi.org/10.1016/j.appet.2004.05.003

Reffstrup, T. K., Larsen, J. C., & Meyer, O. (2010). Risk assessment of mixtures of pesticides. Current approaches and future strategies. *Regulatory Toxicology and Pharmacology*, *56*(2), 174–192. https://doi.org/10.1016/j.yrtph.2009.09.013

Reinert, W. (2015). *Pflanzenschutzmittel – Wirkstoffgenehmigung in der EU*.

Remme, J. H. F., Binka, F., & Nabarro, D. (2001). Toward a framework and indicators for monitoring Roll Back Malaria. *American Journal of Tropical Medicine and Hygiene*, *64*(1-2 SUPPL.), 76–84. https://doi.org/10.4269/AJTMH.2001.64.76

Remme, M., Martinez-Alvarez, M., & Vassall, A. (2017). Cost-Effectiveness Thresholds in Global Health: Taking a Multisectoral Perspective. *Value in Health : The Journal of the International Society for Pharmacoeconomics and Outcomes Research*, *20*(4), 699–704. https://doi.org/10.1016/J.JVAL.2016.11.009

Renn, O. (2015). Stakeholder and public involvement in risk governance. *International Journal of Disaster Risk Science*, *6*(1), 8–20. https://doi.org/10.1007/s13753-015-0037-6

Renn, O., Laubichler, M., Lucas, K., Kröger, W., Schanze, J., Scholz, R. W., & Schweizer, P. (2020). Systemic Risks from Different Perspectives. *Risk Analysis*, risa.13657. https://doi.org/10.1111/risa.13657

Renwick, A. G., Barlow, S. M., Hertz-Picciotto, I., Boobis, A. R., Dybing, E., Edler, L., Eisenbrand, G., Greig, J. B., Kleiner, J., Lambe, J., Müller, D. J. G., Smith, M. R., Tritscher, A., Tuijtelaars, S., Van Den Brandt, P. A., Walker, R., & Kroes, R. (2003). Risk characterisation of chemicals in food and diet. *Food and Chemical Toxicology*, *41*(9), 1211–1271. https://doi.org/10.1016/S0278-6915(03)00064-4

Rhomberg, L. R., Goodman, J. E., Bailey, L. A., Prueitt, R. L., Beck, N. B., Bevan, C., Honeycutt, M., Kaminski, N. E., Paoli, G., Pottenger, L. H., Scherer, R. W., Wise, K. C., & Becker, R. A. (2013). A survey of frameworks for best practices in weight-of-evidence analyses. *Critical Reviews in Toxicology*, *43*(9), 753–784. https://doi.org/10.3109/10408444.2013.832727

Ridder, H. G. (2017). The theory contribution of case study research designs. *Business Research*, *10*(2), 281–305. https://doi.org/10.1007/s40685-017-0045-z

Rieger, J., Weible, D., & Anders, S. (2017). "Why some consumers don't care": Heterogeneity in household responses to

a food scandal. *Appetite*, *113*, 200–214. https://doi.org/10.1016/j.appet.2017.02.038

Robinson, L. A., Hammitt, J. K., Jamison, D. T., & Walker, D. G. (2019). Conducting Benefit-Cost Analysis in Low- and Middle-Income Countries: Introduction to the Special Issue. *Journal of Benefit-Cost Analysis*, *10*(S1), 1–14. https://doi.org/10.1017/BCA.2019.4

Robinson, L. A., Hammitt, J. K., & O'Keeffe, L. (2019). Valuing Mortality Risk Reductions in Global Benefit-Cost Analysis. *Journal of Benefit-Cost Analysis*, *10*(S1), 15–50. https://doi.org/10.1017/BCA.2018.26

Robinson, T., Germini, A., Deluyker, H., Hardy, A., & Liem, D. (2016). Conference conclusions: shaping the future of food safety, together. *EFSA Journal*, *14*(14), 1–9. https://doi.org/10.2903/j.efsa.2016.s0510

Roosen, J., Thiele, S., & Hansen, K. (2005). Food Risk Perceptions by Different Consumer Groups. *Acta Agriculturae Scandinavica, Section C – Food Economics*, *2*(1), 13–26.

Ropeik, D. P. (2011). Risk perception in toxicology-part I: Moving beyond scientific instincts to understand risk perception. *Toxicological Sciences*, *121*(1), 1–6. https://doi.org/10.1093/toxsci/kfr048

Roselius, T. (1971). Consumer Rankings of Risk Reduction Methods. *Journal of Marketing*, *35*(1), 56. https://doi.org/10.2307/1250565

Rosman, L. (1993). Public Participation in International Pesticide Regulation : When the Codex Commission Decides , Who Will Listen ? *Virginia Environmental Law Journal*, *12*(329), 329–365.

Rowe, G., & Frewer, L. J. (2005). A typology of public engagement mechanisms. *Science Technology and Human Values*, *30*(2), 251–290. https://doi.org/10.1177/0162243904271724

Ruzante, J. M., Davidson, V. J., Caswell, J., Fazil, A., Cranfield, J. A. L., Henson, S. J., Anders, S. M., Schmidt, C., & Farber, J. M. (2010). A multifactorial risk prioritization framework for foodborne pathogens. *Risk Analysis*, *30*(5), 724–742. https://doi.org/10.1111/j.1539-6924.2009.01278.x

Saba, A., & Messina, F. (2003). Attitudes towards organic foods and risk/benefit perception associated with pesticides. *Food Quality and Preference*, *14*(8), 637–645. https://doi.org/10.1016/S0950-3293(02)00188-X

Salanié, F., & Treich, N. (2009). Regulation in Happyville. *Economic Journal*, *119*(537), 665–679. https://doi.org/10.1111/j.1468-0297.2009.02221.x

Saltelli, A., & Giampietro, M. (2017). What is wrong with evidence based policy, and how can it be improved? *Futures*, *91*, 62–71. https://doi.org/10.1016/j.futures.2016.11.012

Sapp, S. G., Arnot, C., Fallon, J., Fleck, T., Soorholtz, D., Sutton-Vermeulen, M., & Wilson, J. J. H. (2009). Consumer Trust in the U.S. Food System: An Examination of the Recreancy Theorem. *Rural Sociology*, *74*(4), 525–545. http://content.ebscohost.com/ContentServer.asp?T=P&P=AN&K=46812444&S=R&D=sih&EbscoContent=dGJyMMTo50Seprc4yOvsOLCmr0%2BeprJSrq%2B4SLCWxWXS&ContentCustomer=dGJyMPGvsE2yqLNLuePfgeyx44Dt6fIA

Schlenker, W., & Villas-Boas, S. B. (2009). Consumer and market responses to mad cow disease. *American Journal of Agricultural Economics*, *91*(4), 1140–1152. https://doi.org/10.1111/j.1467-8276.2009.01315.x

Sexton, S. E., Lei, Z., & Zilberman, D. (2007). The economics of pesticides and pest control. *International Review of Environmental and Resource Economics*, *1*(3), 271–326. https://doi.org/10.1561/101.00000007

Shaw, I. (2005). Is it safe to eat? Enjoy eating and minimize food risks. In *Is it safe to eat? Enjoy eating and minimize food risks.* Springer. https://doi.org/10.1007/b1138

Siegrist, M., Cvetkovich, G., & Roth, C. (2000). Salient Value Similarity, Social Trust, and Risk/Benefit Perception. *Risk Analysis: An International Journal*, *20*(3), 353–362. http://search.ebscohost.com/login.aspx?direct=true%7B&%7Ddb=buh%7B&%7DAN=6632602%7B&%7Dsite=ehost-live

Siegrist, M., Gutscher, H., & Earle, T. C. (2005). Perception of risk: The influence of general trust, and general confidence. *Journal of Risk Research*, *8*(2), 145–156. https://doi.org/10.1080/1366987032000105315

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

Sjoberg, L. (2000). Factors in Risk Perception. *Risk Analysis*, *20*(1), 1–11. https://doi.org/10.1111/0272-4332.00001

Skevas, T., Oude Lansink, A. G. J. M., & Stefanou, S. E. (2013). Designing the emerging EU pesticide policy: A literature review. *NJAS - Wageningen Journal of Life Sciences*, *64–65*(65), 95–103. https://doi.org/10.1016/j.njas.2012.09.001

Slikker, W., Andersen, M. E., Bogdanffy, M. S., Bus, J. S., Cohen, S. D., Conolly, R. B., David, R. M., Doerrer, N. G., Dorman, D. C., Gaylor, D. W., Hattis, D., Rogers, J. M., Setzer, R. W., Swenberg, J. A., & Wallace, K. (2004). Dose-dependent transitions in mechanisms of toxicity: Case studies. *Toxicology and Applied Pharmacology*, *201*(3), 226–294. https://doi.org/10.1016/j.taap.2004.06.027

Slimak, M. W., & Dietz, T. (2006). Personal values, beliefs, and ecological risk perception. *Risk Analysis*, *26*(6), 1689–

1705. https://doi.org/10.1111/j.1539-6924.2006.00832.x

Slovic, P. (1987). Perception of risk. *Science*, *236*(4799), 280–285.

Smith-Spangler, C., Brandeau, M. L., Hunter, G. E., Bavinger, J. C., Pearson, M., Eschbach, P. J., Sundaram, V., Liu, H., Schirmer, P., & Stave, C. (2012). Are Organic Foods Safer or Healthier Than Conventional Alternatives? *Annals of Internal Medicine*, *157*, 348–366.

Sparks, P., & Shepherd, R. (1994). Public Perceptions of the Potential Hazards Associated With Food-Production and Food-Consumption - an Empirical-Study. *Risk Analysis*, *14*(5), 799–806. https://doi.org/10.1111/j.1539-6924.1994.tb00291.x

Spiekermann, U. (2011). Redefining food: the standardization of products and production in Europe and the United States, 1880–1914. *History and Technology*, *27*(1), 11–36. https://doi.org/10.1080/07341512.2011.548971

Starr, C. (1969). Social Benefit versus Technological Risk. *Science*, *165*, 1232–1238. http://www.w2agz.com/Library/Chauncey Starr/UCLA & Risk Analysis (1966 - 1971)/(1969) C Starr, Social Benefit versus Technological Risk, Science 165, 1232 (1969).pdf

Statistisches Bundesamt. (2019). *Indikatoren für evidenzbasierte Politik*. https://www.destatis.de/DE/ZahlenFakten/LaenderRegionen/Internationales/Thema/PolitischeIndikatoren.html

Stein, F., & Sridhar, D. (2018). The financialisation of global health. *Wellcome Open Research*, *3*(17). https://doi.org/10.12688/WELLCOMEOPENRES.13885.1

Steketee, R. W., & Campbell, C. C. (2010). Impact of national malaria control scale-up programmes in Africa: Magnitude and attribution of effects. *Malaria Journal*, *9*(299). https://doi.org/10.1186/1475-2875-9-299/FIGURES/10

Stirling, A. (2008). Science, precaution, and the politics of technological risk. Converging implications in evolutionary and social scientific perspectives. *Annals of the New York Academy of Sciences*, *1128*, 95–110. https://doi.org/https://doi.org/10.1196/annals.1399.011

Storck, V., Karpouzas, D. G., & Martin-Laurent, F. (2017). Towards a better pesticide policy for the European Union. *Science of the Total Environment*, *575*, 1027–1033. https://doi.org/10.1016/j.scitotenv.2016.09.167

Stornetta, A., Engeli, B. E., Zarn, J. A., Gremaud, G., & Sturla, S. J. (2015). Development of a risk management tool for prioritizing chemical hazard-food pairs and demonstration for selected mycotoxins. *Regulatory Toxicology and Pharmacology*, *72*(2), 257–265. https://doi.org/10.1016/j.yrtph.2015.04.014

Sunstein, C. R. (2002). *Hazardous Heuristics* (No. 33; Law & Legal Theory Working Paper). https://chicagounbound.uchicago.edu/cgi/viewcontent.cgi?article=1137&context=public_law_and_legal_theory

Sunstein, C. R. (2018). *The cost-benefit revolution*. MIT Press.

Syddell, M. (1990). Anatomy of a disaster. *Agribusiness*, *6*(5), 417–424. https://doi.org/10.3167/foc.2007.490107

Szajkowska, A. (2012). *Regulating food law : risk analysis and the precautionary principle as general principles of EU food law*. Wageningen Academic Publishers. https://books.google.de/books/about/Regulating_Food_Law.html?id=5Xv0hgss2IEC&redir_esc=y

Taebi, B. (2017). Bridging the Gap between Social Acceptance and Ethical Acceptability. *Risk Analysis*, *37*(10), 1817–1827. https://doi.org/10.1111/risa.12734

Teklehaimanot, A., & Mejia, P. (2008). Malaria and poverty. *Annals of the New York Academy of Sciences*, *1136*, 32–37. https://doi.org/10.1196/ANNALS.1425.037

Torraco, R. J. (2005). Writing Integrative Literature Reviews: Guidelines and Examples. *Human Resource Development Review*, *4*(3), 356–367. https://doi.org/10.1177/1534484305278283

Travisi, M. C., Nijkamp, P., & Vindigni, G. (2006). Pesticide risk valuation in empirical economics: A comparative approach. *Ecological Economics*, *56*(4), 455–474. https://doi.org/10.1016/j.ecolecon.2004.06.026

Trumbull, G. (2006). *Consumer capitalism : politics, product markets, and firm strategy in France and Germany*. Cornell University Press.

Tucker, A. J. (2008). Pesticide residues in food - Quantifying risk and protecting the consumer. *Trends in Food Science and Technology*, *19*(SUPPL. 1). https://doi.org/10.1016/j.tifs.2008.07.010

UN. (2020). *United Nations Department of Economic and Social Affairs. Sustainable development goal 2*. https://sdgs.un.org/goals/goal2

Unusan, N. (2007). Consumer food safety knowledge and practices in the home in Turkey. *Food Control*, *18*(1), 45–51. https://doi.org/10.1016/j.foodcont.2005.08.006

Van der Fels-Klerx, H. J., Van Asselt, E. D., Raley, M., Poulsen, M., Korsgaard, H., Bredsdorff, L., Nauta, M., D'agostino, M., Coles, D., Marvin, H. J. P., & Frewer, L. J. (2018). Critical review of methods for risk ranking of food-related hazards, based on risks for human health. In *Critical Reviews in Food Science and Nutrition* (Vol. 58, Issue 2, pp. 178–193). https://doi.org/10.1080/10408398.2016.1141165

van der Meulen, B., & van der Velde, M. (2004). *Food safety law in the European Union*. Wageningen Academic Publishers.

Van Rijswijk, W., & Frewer, L. J. (2008). Consumer perceptions of food quality and safety and their relation to traceability. *British Food Journal*, *110*(10), 1034–1046. https://doi.org/10.1108/00070700810906642//doi.org/10.1108/00070700110386728%22%3Ehttps://

van Schothorst, M. (2002). Microbiological Risk Assessment of foods in international trade. *Safety Science*, *40*(1–4), 359–382. https://doi.org/10.1016/S0925-7535(01)00055-8

van Zwanenberg, P., & Millstone, E. (2000). Beyond skeptical relativism: Evaluating the social constructions of expert risk assessments. *Science Technology and Human Values*, *25*(3), 259–282. https://doi.org/10.1177/016224390002500301

Vanclay, F. (2002). Conceptualising social impacts. *Environmental Impact Assessment Review*, *22*(3), 183–211. https://doi.org/10.1016/S0195-9255(01)00105-6

Vareman, N., & Persson, J. (2010). Why separate risk assessors and risk managers? Further external values affecting the risk assessor qua risk assessor. *Journal of Risk Research*, *13*(5), 687–700. https://doi.org/10.1080/13669871003660759

Verbeke, W., Frewer, L. J., Scholderer, J., & De Brabander, H. F. (2007). Why consumers behave as they do with respect to food safety and risk information. *Analytica Chimica Acta*, *586*(1-2 SPEC. ISS.), 2–7. https://doi.org/10.1016/j.aca.2006.07.065

Verbeke, W., Sioen, I., Pieniak, Z., Van Camp, J., & De Henauw, S. (2005). Consumer perception versus scientific evidence about health benefits and safety risks from fish consumption. *Public Health Nutrition*, *8*(04), 422–429. https://doi.org/10.1079/PHN2004697

Verstraete, F. (2014). *Ensuring safety of foods as regard chemical contaminants through regulation and control*. http://www-pub.iaea.org/iaeameetings/cn222pn/Session5/5-02-IAEA-CN-222-Verstraete-EU.pdf

Vogel, E., & Delfini, G. (2008). The role of evidence in risk characterization : making sense of conflicting data. In P. Wiedemann & H. Schütz (Eds.), *The role of evidence in risk characterization : making sense of conflicting data* (pp. 25–34). Wiley-VCH.

Vos, E. (2000). EU food safety regulation in the aftermath of the BSE crisis. *Journal of Consumer Policy*, *23*(3), 227–255. https://doi.org/10.1023/A:1007123502914

Wagner, M. (2016). Law Talk v. Science Talk: The Languages of Law and Science in WTO Proceedings. *Fordham International Law Journal*, *35*(1), 151-. http://ir.lawnet.fordham.edu/ilj

Walker, V. R. (2001). Consistent Levels of Protection in International Trade Disputes: Using Risk Perception Research to Justify Different Levels of Acceptable Risk. *Environmental Law Reporter*, *31*, 11317. http://people.hofstra.edu/faculty/vern_r_walker/Walker2.pdf

Weed, D. L. (2005). Weight of evidence: A review of concept and methods. *Risk Analysis*, *25*(6), 1545–1557. https://doi.org/10.1111/j.1539-6924.2005.00699.x

Wentholt, M. T. A., Rowe, G., König, A., Marvin, H. J. P., & Frewer, L. J. (2009). The views of key stakeholders on an evolving food risk governance framework: Results from a Delphi study. *Food Policy*, *34*(6), 539–548. https://doi.org/10.1016/j.foodpol.2009.06.002

Whaley, P., & Halsall, C. (2016). Assuring high-quality evidence reviews for chemical risk assessment: Five lessons from guest editing the first environmental health journal special issue dedicated to systematic review. *Environment International*, *92–93*, 553–555. https://doi.org/10.1016/j.envint.2016.04.016

WHO. (2015). *WHO estimates of the global burden of foodborne diseases*. https://doi.org/10.1007/978-3-642-27769-6_3884-1

Wilcock, A., Pun, M., Khanona, J., & Aung, M. (2004). Consumer attitudes, knowledge and behaviour: A review of food safety issues. In *Trends in Food Science and Technology* (Vol. 15, Issue 2, pp. 56–66). https://doi.org/10.1016/j.tifs.2003.08.004

Williams, P. R. D., & Hammitt, J. K. (2001). Perceived Risks of Conventional and Organic Produce: Pesticides, Pathogens, and Natural Toxins. *Risk Analysis*, *21*(2), 319–330. https://doi.org/10.1111/0272-4332.212114

Wilson, J. S., & Otsuki, T. (2004). To spray or not to spray: Pesticides, banana exports, and food safety. *Food Policy*, *29*(2), 131–146. https://doi.org/10.1016/j.foodpol.2004.02.003

Wilson, R. S., Zwickle, A., & Walpole, H. (2019). Developing a Broadly Applicable Measure of Risk Perception. *Risk Analysis*, *39*(4), 777–791. https://doi.org/10.1111/risa.13207

WTO. (1995). *Agreement on the Application of Sanitary and Phytosanitary Measures*. https://doi.org/doi:10.1163/ej.9789004145641.i-565.41

WTO. (2021a). *The WTO and the FAO/WHO Codex Alimentarius*. WTO Webpage.

https://www.wto.org/english/thewto_e/coher_e/wto_codex_e.htm

WTO. (2021b). *The WTO and the International Plant Protection Convention (IPPC)*. WTO Webpage. https://www.wto.org/english/thewto_e/coher_e/wto_ippc_e.htm

Yamey, G. (2001). Global campaign to eradicate malaria : Roll Back Malaria has achieved a high profile but little real action. *BMJ : British Medical Journal*, *322*(7296), 1192. https://doi.org/10.1136/BMJ.322.7296.1191

Yeung, R. M. W., & Morris, J. (2001). Food safety risk. Consumer perception and purchase behaviour. *British Food Journal*, *103*(3), 170–186. http://www-1emeraldinsight-1com-1005a77i50069.emedia1.bsb-muenchen.de/doi/pdfplus/10.1108/00070700110386728

Yeung, R. M. w., & Yee, W. M. s. (2002). Multi-dimensional analysis of consumer-perceived risk in chicken meat. *Nutrition & Food Science*, *32*(6), 219–226. https://doi.org/10.1108/00346650210454208

Yin, R. K. (2013). *Case study research : design and methods*. SAGE.

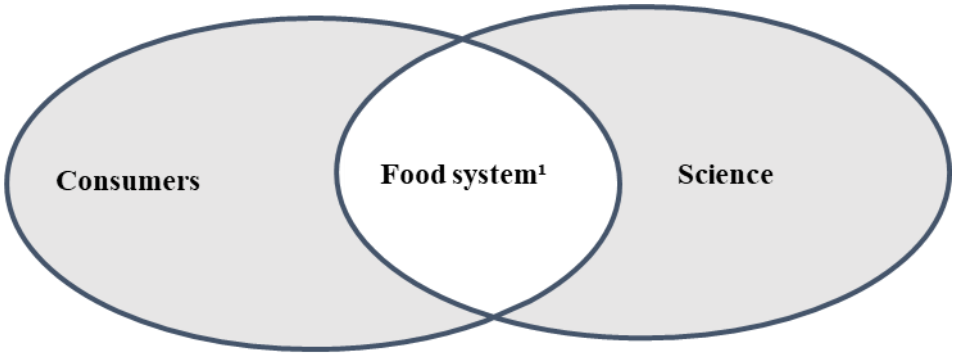
Zachmann, K. (2014). Risk in historical perspective: Concepts, contexts, and conjunctions. In *Risk - A Multidisciplinary Introduction* (pp. 3–35). Springer International Publishing. https://doi.org/10.1007/978-3-319-04486-6_1

Zachmann, K., & Ehlers, S. (2019). Wissen und Begründen – Evidenz als umkämpfte Ressource in der Wissensgesellschaft. In *Practicing Evidence - Evidencing Practice*.

Zwietering, M. H. (2015). Risk assessment and risk management for safe foods: Assessment needs inclusion of variability and uncertainty, management needs discrete decisions. *International Journal of Food Microbiology*, *213*, 118–123. https://doi.org/10.1016/j.ijfoodmicro.2015.03.032

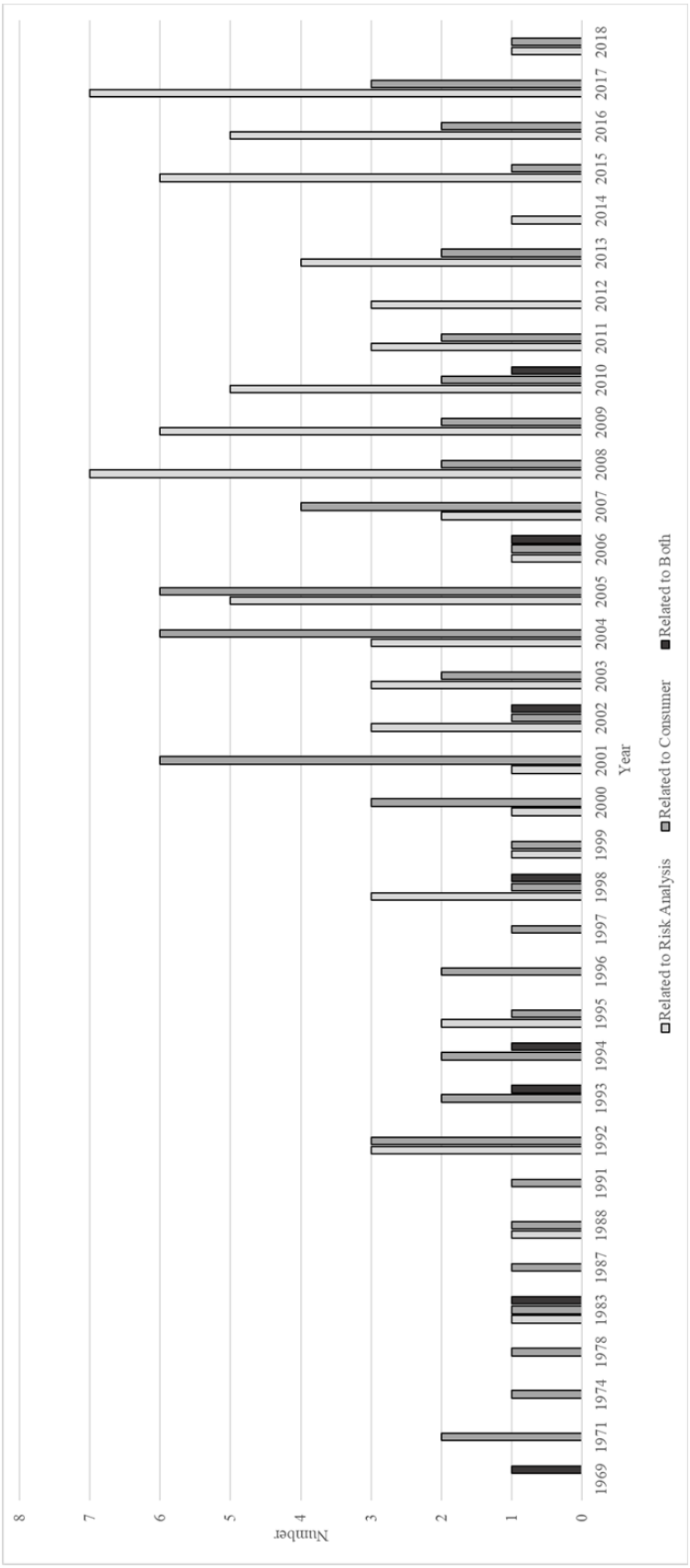
6. Appendices

Appendix A to Section 3.1. Guiding Theory



[1] The food system includes all activities in producing food, processing food, packaging and distributing food, and retailing and food consumption. A variety of actors controls and participates in these activities (for a more detailed description s. Ingram (2011)).

Appendix B to Section 3.1. Overview literature



Appendix C to Section 3.2. Interview guide

Risk assessment

Main	Sub	Question type Patton
1. Please describe your personal background in the field in a few words.	- Years in risk assessment - Personal expertise	Past x Background
At first we will focus on the consumer-related side, in particular the consumers respectively the public		
2. First, we would like to discuss the relevance of consumer-related information in risk assessment.		
	2.1. Which consumer-related data and studies do you consider if you conduct a risk assessment? - Standard operations? Ad-Hoc? - Where in the procedure involved?	Present x Experience/Knowledge
Note: "Consumer-related data and studies" is chosen as a very broad term on purpose. It might refer to exposition or consumption but also acceptance, trust and/or public crises.		
	2.2. Do you remember a particular case from the past where further, atypical consumer-related data and studies were considered in the risk assessment?	
	2.3. From which sources do you usually consider consumer-related data? - Which protagonist for the production of evidence?	Past/Present x Knowledge
	2.4. Internally-produced/commissioned studies: Which quality criteria need to be fulfilled by these studies? How are these quality criteria codified in XXX documents?	Present x Knowledge
Note: Quality criteria do not necessarily need to be "classical" scientific quality criteria but can include usability, efficiency, costs and/or other criteria which are relevant for you.		
	2.5. Externally-produced studies: Which quality criteria need to be fulfilled by these studies? How do you decide about in- or exclusion (of) these studies? How are these quality criteria codified in EFSA documents?	Present x Knowledge
Note: Quality criteria do not necessarily need to be "classical" scientific quality criteria but can include usability, efficiency, costs and/or other criteria which are relevant for you.		
	2.6. In your opinion, what are some major issues in the collection and use of consumer-related data and studies? - In what way problematic and how does it affect risk assessment?	Present x Opinion/Values
3. **We also observe influential tendencies in food safety through consumers as, for example, participation procedures or initiatives.**		
	3.1. Which types of influence did you observe in your area of risk assessment? - Type - Form - Tendency	Past x Knowledge
	3.2. How did these influences affect your work in risk assessment? - Tendency	Past x Behavior/Experiences
	3.3.	Present x Opinions/Values
Now we have look on supply related side.		

4. Next, we would like to discuss the relevance of supply-related information in risk assessment.	Repetition part two	Repetition part two
5. We also observe influential tendencies on risk assessment through supply side as, for example, international trade agreements or private-sector standards.	Repetition part three	Repetition part three
6. Which current societal developments will change the practice of risk assessment in your field in the future?		Future x Opinion/Values
	6.1. How will these developments change the current practice of risk assessments?	Future x Opinion/Values
7. Would you like to add more aspects we have not discussed yet?		

Risk communication

Main	Sub	Question type Patton
1. Please describe your personal background in the field in a few words.	- Years in risk assessment - Personal expertise	Past x Background
2. First, we would like to discuss the relevance of consumer-related information in risk communication/stakeholder engagement.		Past x Background
	2.1. With which type of consumer-related data and studies or information are you dealing in your work in risk communication/stakeholder engagement?	Present
Note: "Consumer-related data and studies" is chosen as a very broad term on purpose. It might refer to exposition or consumption but also acceptance, trust and/or public crises.		
	2.2. Do you remember a particular case from the past where further, atypical consumer-related data and studies were considered in your work? - Special case - Is there any reason, why these data/surveys were used in this case?	Past x Experience
	2.3. From which sources do you usually consider consumer-related data? - Which stakeholders for the production of evidence?	Past/Present x Knowledge
	2.4. Internally-produced/commissioned studies: Which quality criteria need to be fulfilled by these studies? How are these quality criteria codified in XXX documents?	Present x Knowledge
Note: Quality criteria do not necessarily need to be "classical" scientific quality criteria but can include usability, efficiency, costs and/or other criteria which are relevant for you.		
	2.5. Externally-produced studies: Which quality criteria need to be fulfilled by these studies? How do you decide about in- or exclusion these studies? How are these quality criteria codified in XXX documents?	Present x Knowledge
	2.6. In your opinion, what are some major issues in the collection and use of consumer-related data and studies in your work? - In what way problematic and how does it affect risk communication?	Present x Opinion/Values

"On the other hand, we also see that consumer increasingly influence the risk communication"		
3. We also observe influential tendencies in food safety through consumers as, for example, participation procedures or initiatives.		Present x Experience/ Knowledge
	3.2. Which types of influence did you observe or include in work? - Type - Form - Tendency	Past x Experience/ Knowledge
	3.3. Which differences did you observe between different types of risks (e.g., microbiological risks vs. pesticide residues?	Past x Knowledge
	3.4. How did these influences affect your work? - Tendency	Past x Behavior/ Experience
4. Next we would like to discuss the relevance of supply-related information in risk communication/stakeholder engagement.	Repetition part two	Repetition part two
6. We also observe influential tendencies through supply side as, for example, international trade agreements or private-sector standards.	Repetition part three	Repetition part three
7. How do you interact with risk assessment in your work in terms of data and information exchange?		
8. Which current societal developments will change the practice of risk communication in your field in the future?		Future x Opinion/Values
	8.1. How will these developments change the current practice of risk communication?	Future x Opinion/Values
9. Would you like to add more aspects we have not discussed yet?		x

Appendix D to Section 3.2. Detailed presentation of themes and sub-themes

Input

	Risk assessment		Risk communication
	Consumer	**Food sector**	**Consumer**
Collect data* (Vertical)	Surveys Databases Literature	Literature Databases Model/Lab studies Monitoring Targeted consultations (M) Discussions	News screening Databases Surveys Expert studies Literature Consumer conferences Focus groups Combinations
Provide input* (Vertical)	Public Consultations (P) Call for data Crowd-sourcing/Citizen science Only scientific input (P)	Self-monitoring	
Quality criteria input/ Filter 1* (Horizontal)	**External expectations** Independence Transparency **Context** Data pool depends on question/context Feasibility and costs Data selection (partly) regulated Stricter than RC **"Classic" scientific quality criteria** Scientific standards/peer-review International Standards Consideration of limitations/uncertainty (M) Representativety/Comparability/Harmonization **Risk-assessment expertise** Own criteria/approach, Weight of Evidence Expertise/Control/Review RA External providers assessed, supervised, monitored		**Context** Data pool depends on question/context Efficiency **"Classic" scientific quality criteria** Standardization Plausibility checks No defined quality criteria, based on Cochrane **Risk communication expertise** External providers assessed, supervised, monitored Defined process chains
***In risk assessment columns: If sub-themes are only mentioned in one case they are marked with (M) - microbiological hazards or (P) - plant protection products. If sub-themes occurred in both cases, they are not marked.			

Processing

	Risk assessment		Risk Communication
	Consumer	**Food sector**	**Consumer**
Include information* (Vertical)	Trends (indirectly) Knowledge Exposition Behavior Consumption Habits/Practices (M) Toxicity** (P) Interaction Matrix-Consumer-Pathogen** (M)	Production hygiene (M) Prevalence (M) Safety/efficiency (M) Processing factors Residues/Occurrence Product volumes/distribution patterns (M) Supply data no role for regulated products (P)	Perception/Concerns Trends Knowledge Information Sources Acceptance/Trust
Influence* (Vertical)	**Types of influence** Influence relevant questions (M) RA needs to consider society, Consideration can damage independence (M) Generate discourse (P) Initiate data collection (P) Initiate method development (P) Initiate change in RA (P) Do not initiate change in RA Binding capacities in RA Influence amount of work, not content (P) Request independent studies **Evaluation by experts** positive, neutral, negative **Instruments** Dialogue events, Citizen portal, Citizen hotline, Freiheitsinformationsgesetz	**Types of influence** Try to use RA to increase visibility (M) Rarely, obvious, not influential (M) Mainly on RM level (P) Secondary safety standards thwart primary safety standards (P)	Initiate change in communication
Involve (Vertical)	**Types of involvement** Enable two-way communication, interaction Interaction with stakeholder representatives in whole RA process Objective: Establish trust, based on core values Influence difficult to measure Output considered in RA/RC **Quality criteria** Reputation barometers as assessment tools Defined processes and criteria, transparent **Instruments** Discussion groups on emerging risks, ad-hoc groups Working groups Stakeholder bureaus Annual gatherings		
Interaction risk communication → risk assessment (Horizontal)	Risk screenings as feedback/source for RA Raising awareness Request/Receive information Contextualize/ Reduce/Translate information Relevant topics initiated by RA Intensive/important exchange Tasks separated legally		

***** In risk assessment columns: If sub-themes are only mentioned in one case they are marked with (M) - microbiological hazards or (P) - plant protection products. If sub-themes occurred in both cases, they are not marked.
**Does not fall into consumer-related data explicitly under our definition (more: hazard properties) but was reported by some experts, as it considers toxicological health endpoints in consumers/interaction of pathogen with consumer.

Output

Publish/ Communicate* (Vertical)	**Forms** Risk assessments Revisions of risk assessments Scientific publications Reports Scientific Opinions Risk screenings Communication tools/Pilot tools Presentation tools Social Media Traditional publication formats Campaigns **Issues** Scientific facts in communication often not important Challenge to communicate RA and scientific content Challenge to communicate RA without RM action
Communicate via outside communicators (Media)* (Vertical)	Close interaction Analyzed in media analysis Invited for opinion exchanges Protected by freedom of press Receive clarification letters Attacks RA Communicate driven by sales Pose requests, generate work Generate public requests/fears/interest Distribute fake news Amplify and manipulate Interpret and bias information
Communicate via outside Communicators (Multiplicators)* (Vertical)	Exchange Stakeholder experts as testing ground for RC Used for communication by RC, direct line to consumers
Interact * (Vertical)	Interaction depends on information channel Measurement of engagement More influence from RA to consumer than other way round Respond mechanisms: for example, scientist empowerment Validation of communication prototypes
Filter 2* (Horizontal)	Output depends on role/level (Member Sate/European) Content and form (partly) determined by regulation Type/Form depends on question/context/availability Independence/Transparency/Harmonization Specific quality criteria for scientific opinions Publishing standards Communication models Cooperation with risk management
* Outputs are not directly relatable to risk assessment or communication because often joint effort.	

Context

Context*	**External** Public mistrusts public organizations generally Public discourse relates to non-risk assessment topics Other actors influence acceptance of facts Trade/Globalization impacts due to imported pests Safety debates in scientific community influence public discourse Policy poses requests Policy makes accusations Interest groups criticize, initiate change Interest groups stimulate/sensitize public Interest groups research and publish Crisis/issues initiate change in risk assessment (rarely) **Risk analysis** Risk management responsible for framework, law, related communication Superordinate organizations initiate change in risk assessment Constant revision risk-assessment practices Risk assessment develops with scientific progress Risk assessment initiates development of risk assessment Risk assessment cooperates with independent scientific experts Risk assessment participates in international RA organizations Cooperation national/European level, different roles
* Context not directly relatable to risk assessment or communication, therefore it describes the context of the risk assessment organization.	