

Sustainable Consumption – Insights from Text Mining on Consumer Perception and Agenda Setting and Implications for Consumer Behavior

Hannah Danner

Vollständiger Abdruck der von der Fakultät für Wirtschaftswissenschaften der Technischen Universität München zur Erlangung des akademischen Grades eines Doktors der Wirtschaftswissenschaften (Dr. rer. pol.) genehmigten Dissertation.

Vorsitzender: Prof. Dr. Jörg Königstorfer
Prüferinnen der Dissertation: Prof. Luisa Menapace, Ph.D.
Prof. Dr. Jutta Roosen

Die Dissertation wurde am 07.05.2021 bei der Technischen Universität München eingereicht und durch die Fakultät für Wirtschaftswissenschaften am 15.07.2021 angenommen.

Acknowledgements

First and foremost, I thank Prof. Dr. Jutta Roosen and Prof. Dr. Luisa Menapace for the supervision of my doctoral thesis and for providing an open and supportive environment to develop and implement my research ideas and projects. Further thanks go to Prof. Dr. John Thøgersen for receiving me as a visiting scholar at Aarhus University, the research collaboration, and his advice throughout my dissertation process. Thank you to Prof. Dr. Jörg Königstorfer for heading the dissertation commission.

I thank *Studienstiftung des deutschen Volkes* for the scholarships and the inspirational environment provided throughout my Bachelor, Master, and Ph.D. studies.

Thanks for the great company to my colleagues at the Chair of Marketing and Consumer Research.

Thank you to my family and friends who were with me throughout the last years and especially for the unconditional support of my partner.

Summary

Detrimental effects of food production and consumption on the environment become more and more evident and lead to the rise of sustainability-differentiated product categories, most prominently organic food. To spur sustainable consumption, understanding consumer concern and behavior is key. Consumers increasingly voice their opinions on online platforms like discussion boards and news websites. This dissertation exploits such user-generated content to advance the understanding of consumer perception and behavior regarding the example of organic food. In an interdisciplinary, mixed-method research approach, different text-mining techniques were applied and survey data including a choice experiment were used to explore the behavioral relevance of what consumers discuss online.

The findings of this research show the diversity of beliefs and topics discussed online and how salient they are to consumers in the German-speaking and US contexts, which are the largest organic food markets globally. Further, the news media were found to be important agenda setters influencing which topics are salient among their readers. Finally, the results of this thesis confirmed that pro-organic judgment and decision-making can be boosted by priming topics that are salient in online environments.

This work contributes to consumer research regarding consumer perception and behavior with respect to organic food. It demonstrates the potential of user-generated content and employed both qualitative and novel machine-learning-based text-mining techniques. The thesis provides insights and implications for research and practice regarding the online salience of organic food beliefs and topics, the media as an influential determinant, and how online salience can be applied for effective marketing and policy communication.

List of contents

Summary	III
List of contents	V
List of Tables	VIII
List of Figures	IX
List of Appendices	X
List of Abbreviations	XI
1 Introduction	1
1.1 Toward more sustainable consumption – The case of organic food	1
1.2 Consumer perception of organic food.....	3
1.3 Social media and user-generated content.....	4
1.4 Research aim	7
1.5 Structure of the dissertation	9
2 Background	11
2.1 The four perspectives of the dissertation.....	11
2.1.1 Organic food beliefs in online comments – A qualitative perspective (Essay I)	11
2.1.2 Organic food beliefs in online comments – Combining text-mining techniques (Essay II)	12
2.1.3 Agenda setting in organic food (Essay III)	13
2.1.4 Priming organic food behavior based on online salience (Essay IV)	14
2.2 Organic food markets and organic food consumption	15
2.3 Research approach	18
2.3.1 Overview of data collection and analysis.....	18
2.3.2 Text mining	20
2.3.3 Theories and measures of consumer perception and behavior	24
2.3.3.1 Beliefs and attitudes	24
2.3.3.2 Choices.....	27
3 Four essays	29
3.1 Using online comments to explore consumer beliefs regarding organic food in German- speaking countries and the United States (Essay I).....	29
3.1.1 Introduction.....	30
3.1.2 Review of literature about organic food beliefs	32
3.1.3 Material and methods.....	33
3.1.3.1 Data collection	33
3.1.3.2 Data sampling.....	35
3.1.3.3 Data analysis	36

3.1.4	Results.....	38
3.1.4.1	Beliefs regarding the product.....	41
3.1.4.2	Beliefs regarding the food system.....	42
3.1.4.3	Beliefs regarding authenticity.....	45
3.1.4.4	Beliefs regarding production.....	48
3.1.5	Discussion.....	49
3.1.5.1	Reflections on the method.....	50
3.1.5.2	Research implications.....	51
3.1.6	Conclusions.....	54
3.2	Combining content analysis and neural networks to analyze discussion topics in online comments about organic food (<i>Essay II</i>).....	55
3.2.1	Introduction.....	56
3.2.2	Methodology.....	57
3.2.2.1	Step 1 – Qualitative analysis.....	57
3.2.2.2	Step 2 – Universal Sentence Encoder.....	57
3.2.2.3	Step 3 – Evaluation.....	58
3.2.3	Results.....	59
3.2.4	Discussion.....	60
3.2.5	Conclusion.....	61
3.3	The news media and their audience: agenda setting on organic food in the United States and Germany (<i>Essay III</i>).....	63
3.3.1	Introduction.....	64
3.3.2	Theoretical background.....	66
3.3.3	Methodology.....	68
3.3.3.1	Data.....	68
3.3.3.2	Topic modeling.....	68
3.3.3.3	Agenda-setting metrics.....	72
3.3.4	Results.....	73
3.3.4.1	The topics.....	73
3.3.4.2	Synchronous topic correlations.....	75
3.3.4.3	Cross-lagged topic correlations.....	76
3.3.4.4	Agenda diversity.....	78
3.3.4.5	Agenda distance.....	81
3.3.5	Discussion.....	83
3.3.6	Conclusions.....	85
3.4	Does online chatter matter for consumer behavior? A priming experiment on organic food (<i>Essay IV</i>).....	87
3.4.1	Introduction.....	88

3.4.2	Theoretical background.....	90
3.4.3	Text-mining study on the online salience of organic food topics	94
3.4.3.1	Method	95
3.4.3.2	Results.....	95
3.4.3.3	Selection of organic food topics for the priming experiment.....	96
3.4.4	Priming experiment.....	97
3.4.4.1	Methods.....	98
3.4.4.1.1	Priming.....	98
3.4.4.1.2	Procedure	99
3.4.4.1.3	Choice experiment design	99
3.4.4.2	Results	100
3.4.4.2.1	Choice experiment	101
3.4.4.2.2	Policy acceptance.....	104
3.4.5	Discussion	105
3.4.6	Conclusion and implications	107
4	Conclusion.....	110
4.1	Main findings and discussion.....	110
4.2	Research implications	118
4.3	Managerial and policy implications	121
	References.....	126
	Appendices.....	149

List of Tables

Table 1 Overview of essays and their research questions, methodology, key findings, contributions, and triangulation.	10
Table 2 Pool of relevant sources.	35
Table 3 The category system of organic food beliefs.	39
Table 4 A seed sentence from content analysis and the 11 sentences identified as similar by USE (cosine similarity = 0.79).	60
Table 5 The most representative words (top words) for each topic in the US media and public agendas.	72
Table 6 The most representative words (top words) for each topic in the German media and public agendas.	72
Table 7 The four agenda-setting metrics and how they measure the relationship between the media and public agendas.	73
Table 8 Topic correlations ρ_i in media and public agenda for the US and Germany.	76
Table 9 Cross-lagged correlations between articles in $t1$ and comments in $t2$ (P_{X1Y2}), and between comments in $t1$ and articles in $t2$ (P_{Y1X2}).	78
Table 10 Latent class multinomial logit analyses for the choices of eggs and milk for the four priming conditions (weak/strong animal welfare (AW), weak/strong biodiversity (BD)) and the control group (CG) moderated by attitude strength and experience with buying organic food.	103
Table 11 ANCOVA for policy acceptance moderated by attitude strength.	105

List of Figures

Figure 1 Organic retail sales share (%) from 2002 to 2018 in Germany and the USA.	16
Figure 2 Pearson correlation of class frequencies (21 topic classes) between content analysis and USE.	59
Figure 3 Relative class frequencies (21 topic classes) in content analysis and USE.	60
Figure 4 Sentence embeddings.	70
Figure 5 Topic modeling based on clustering of pre-trained sentence embeddings.	70
Figure 6 Average topic distributions (%) in US and GER media and public agenda.	75
Figure 7 Total cross-lagged correlations.	78
Figure 8 Agenda diversity over time.	79
Figure 9 Topic distributions.	82
Figure 10 Agenda distance.	83
Figure 11 Research framework and hypotheses.	94
Figure 12 Overall share of comments on organic food mentioning each topic (2007-2017) in <i>spiegel.de</i>	96
Figure 13 Share of comments on organic food mentioning the topics <i>animal welfare</i> and <i>biodiversity</i> over time (2007-2017) in <i>spiegel.de</i>	97
Figure 14 Examples of the eggs and milk choice sets.	100

List of Appendices

Appendix 1 to Section 3.3 Synchronous Topic Correlations.	149
Appendix 2 to Section 3.3 Cross-Lagged Topic Correlations.	149
Appendix 3 to Section 3.3 Normalized Shannon Entropy.	149
Appendix 4 to Section 3.3 Agenda Distance - Jensen-Shannon Distance (JSD).	150
Appendix 5 to Section 3.4 Sample characteristics.....	151
Appendix 6 to Section 3.4 Means (M) and standard deviations (SD) for all self-reported variables (N = 1118).	152
Appendix 7 to Section 3.4 Priming items.....	153

List of Abbreviations

AW	animal welfare
BD	biodiversity
CG	control group
DACH	Germany, Austria, and Switzerland
e.g.	exempli gratia
EU	European Union
GMO	genetically modified organism
i.e.	id est
JSD	Jensen-Shannon Distance
LDA	Latent Dirichlet Allocation
LIWC	Linguistic Inquiry and Word Count
NLP	Natural Language Processing
RCB	Rozelle-Campbell Baseline
UGC	user-generated content
US/USA	United States of America
USDA	United States Department of Agriculture
USE	Universal Sentence Encoder
w.r.t.	with respect to

1 Introduction

1.1 Toward more sustainable consumption – The case of organic food

With their use of natural resources, current consumption levels and patterns—especially those of societies of the Global North—challenge the planetary boundaries. They describe the environmental limits of the earth system within which humans can live safely and societal welfare is ensured (Steffen et al., 2015). Therefore, sustainable consumption has been set as one out of the 17 goals on the *Agenda 2030 for Sustainable Development* (United Nations, 2020). Sustainable consumption refers to the consumption of products that make a contribution to the aims of sustainable development, which are of economic, social and environmental kind (World Bank, 2003). Food consumption is responsible for a large share of the environmental impact of households (Tukker et al., 2010). Worldwide, food supply chains are responsible for almost a quarter of global anthropogenic greenhouse gas emissions (Poore & Nemecek, 2018). Further environmental impacts of food production and consumption include water pollution, soil depletion and biodiversity loss (Reisch et al., 2013; Tsiafouli et al., 2015); other issues are land-use conflicts and increasing health costs (Reisch et al., 2013).

One suggested strategy toward more sustainable food consumption is organic agriculture (Müller et al., 2017). Diets composed of organic food have been found to have a significantly lower environmental impact than those based on conventionally-produced food (Baroni et al., 2007). The environmental benefits of organic farming lie particularly in the ban of synthetic pesticides and fertilizers, which leads to improved nitrogen and phosphorus management, reduced ecotoxicity and eutrophication (and thus the protection of biodiversity), and reduced fossil energy use (Baroni et al., 2007; Müller et al., 2017; Tuck et al., 2014). However, in view of global food security, conversion to organic farming is also controversially discussed as yields are lower compared to conventional agriculture, yield differences vary from 5 to 34% depending on system and local

characteristics according to a meta-analysis (Seufert et al., 2012). Accounting for the resulting increased land use of organic farming, a food systems model by Müller et al. (2017) revealed that a full conversion to organic farming is feasible if accompanied by reductions of food loss and wastage and consumption of animal products, as feed production competes with food production for arable land. Further benefits of organic farming and food have been discussed. Evidence on organic farming being better adapted to climate change than conventional agriculture is yet scarce (Müller et al., 2017). Furthermore, the consumption of organic food has been found to have certain positive effects on human health, especially due to reduced pesticide and antibiotics use in organic farming (Mie et al., 2017; Smith-Spangler et al., 2012). Expected health benefits are thus a main motive for consumers to buy organic food (Federal Office for Agriculture and Food, 2020; Rana & Paul, 2020). However, meta-analytical research has shown that organic food is unlikely to be more nutritious compared to conventional food (Smith-Spangler et al., 2012). An ethical merit of organic over conventional food production lies in increased animal welfare levels established by organic standards on animal husbandry (Spoolder, 2007; Tuck et al., 2014).

In sum, organic food has emerged as a product category, which is currently seen as the most successful among the categories differentiated by sustainability (Jackson et al., 2020; Willer et al., 2019). It is by far the most common sustainability standard with almost 70 million hectares of agricultural area cultivated under organic regulation worldwide corresponding to 1% of the global agricultural land (Willer et al., 2019). In the European Union (EU), 7.5% of the utilized agricultural area are currently under organic production (EUROSTAT, 2020a). Recognizing the contribution of organic farming to environmental protection and the reversing of environmental degradation, the European Green Deal from 2019 sets out the target to reach a share of 25% of organically farmed land by 2030. This goal has also been adopted by the Common Agricultural Policy of the EU and its *Farm to Fork* and *Biodiversity* strategies to shift toward environmentally-friendly food systems and healthy nutrition (European Commission, 2020b). To conclude, due to the outlined sustainability merits and its proliferation, organic food is deemed a widespread example of sustainable consumption.

1.2 Consumer perception of organic food

Consumers are key players in the transition toward more sustainable food systems (Vittersø & Tangeland, 2015). They are both—part of the problem of current food systems and part of the solution of more sustainable food systems (Klößner, 2013). Not only research and policy makers, but also consumers perceive organic food as a sustainable alternative (Siegrist et al., 2015; Verain et al., 2015). With their purchases, consumers can contribute to the development of sustainable food systems, and, in particular, the growth of the organic sector. However, the relative share of organic food sales compared to total food sales remain low, despite the increasing availability of organic food products, consumers' familiarity with the label, and decreasing prices (Chekima et al., 2017; Janssen, 2018). Therefore, further research is needed: Understanding what “organic” means to consumers and how they make decisions is vital to spur demand for organic food (Klößner, 2013). It is crucial to map consumers' current perception of organic food as an important antecedent of organic food purchases.

The controversies regarding the benefits of organic farming outlined in the previous section are also reflected in the perception of consumers, who controversially discuss the advantages and disadvantages of organic food (Olson, 2017). Hence, despite national and supra-national standards setting the criteria to define organic food, consumers have shown to have a diverse understanding of organic food (Lee & Hwang, 2016). Consumer perception of organic food is heterogeneous as organic labeling entails several credence attributes (Lee & Hwang, 2016). As opposed to search (e.g., price, appearance) and experience attributes (e.g., taste), consumers cannot verify credence attributes (e.g., health benefits, environmental impact, production practices) neither before nor after consumption (Darby & Karni, 1973; Nelson, 1970). Previous research has found that search, experience, and credence attributes explain organic food behavior (Fernqvist & Ekelund, 2014; Gracia & Magistris, 2008; Lee & Hwang, 2016). However, credence attributes are of particular importance: In a meta-analysis of studies on organic food consumption published between 1991 and 2016, Massey et al. (2018) have found that consumers value credence attributes more than search and experience attributes. The organic food market is characterized by asymmetric information as consumers are bound to rely on labeling (Massey et al., 2018). At the same time, confusion on what organic

labeling actually implies reigns among consumers. The term “organic” is often not clearly distinguished from or perceived as similar to terms like “natural”, “green”, “sustainable”, “local”, or “fair trade” (Bezawada & Pauwels, 2013; Harper & Makatouni, 2002; Jensen et al., 2019; Vega-Zamora et al., 2013). Therefore, consumers’ choices rely on what they believe organic is or is not.

Consumers’ heterogeneous understanding and critical discussion of organic food make consumer perception a relevant study subject. To improve the understanding of consumer perception and behavior regarding organic food, this dissertation explores a little exploited data source: user-generated content, that is what consumers write on online discussion platforms. The following section outlines the potential of user-generated content to provide detailed insights into consumer rationales concerning organic food and regarding the promotion of this emerging sustainability-differentiated food product category.

1.3 Social media and user-generated content

Worldwide, internet usage experiences a steady increase with users spending a daily average of 2.5 hours online in 2018, a figure expected to rise to over 3 hours by 2021 (Zenith, 2019). Today, 95% of Germans have access to the Internet (EUROSTAT, 2020b), 48% use social networks such as Facebook or Twitter, 30% read news online, and 20% write comments in blogs, discussion boards, or news websites (IfD Allensbach, 2020). For sustainable consumption, digitization and online consumer behavior offer both risks and opportunities (Börjesson Rivera et al., 2014). In addition to the immediate effects of the production, use, and disposal of information and communication technology (e.g., devices), internet usage may spur both sustainable and unsustainable lifestyles. Online shopping environments can facilitate overconsumption, but can also grant access to more sustainable alternatives as well as social media can provide platforms to raise awareness to sustainable consumption (Börjesson Rivera et al., 2014; Frick & Matthies, 2020; Midden et al., 2007). Consumer perception and consumption motives may be influenced by the online sphere, for example, via (news) media consumption, peer-to-peer content and communication, or advertising (Cooper et al., 2012; Frick et al., 2020; Thøgersen, 2014). Frick et al. (2020) have shown that online environments (i.e., the

exposure to online advertisement and social media peer content) influence consumption motives and levels. Thus, research into sustainable consumption cannot ignore the online sphere, as consumption patterns and their sustainability are increasingly discussed in and influenced by online environments (Chatzidakis & Mitussis, 2007; Reisch, 2001).

Consumers share their opinions more and more online on social media (Netzer et al., 2012; Ziegele et al., 2017). Social media refer to online services that enable users to communicate, collaborate, and connect (Sloan & Quan-Haase, 2017). These platforms can take different forms: Social networks such as Facebook and Twitter, media-sharing sites such as YouTube, discussion boards such as Quora, user or company-sponsored blogs and discussion boards, product review boards on e-commerce platforms, or collaborative websites such as Wikipedia (Mangold & Faulds, 2009). Social media facilitate communication between different actors, such as among users, between businesses and their customers, or between news websites and their reader audience (Berger et al., 2020; Sloan & Quan-Haase, 2017). Along with its digitization, journalism has become more participatory as comment boards under news articles provide feedback opportunities to readers, turning them into senders (Santana, 2011).

Content shared by users online, that is user-generated content (UGC), is the essence of social media; it comprises textual, visual, and audio data (Balducci & Marinova, 2018; Tang et al., 2014). This thesis employs textual UGC (hereafter UGC). UGC is extant document data and as such of inherent qualitative nature (Bowen, 2009). Synonymous or similar terms for textual UGC are, e.g., unstructured data (Balducci & Marinova, 2018), social media data (Vidal et al., 2018), social media peer content (Frick et al., 2020), online chatter (Tirunillai & Tellis, 2012), or electronic word-of-mouth (Babić Rosario et al., 2016; Hennig-Thurau et al., 2004). Online reviews (Mudambi & Schuff, 2010) and online recommendations (Cascio et al., 2015) are further sub-types of UGC (for conceptualization of UGC see also, Zablocki et al., 2019).

UGC comes with several characteristics that make it an insightful alternative or complement to traditional data from surveys, experiments, and interviews (Bowen, 2009). While UGC can be used to verify evidence from other sources, it can also uncover what remains unsaid in qualitative data from interviews or focus groups. Newly explored concepts can then be incorporated in subsequent confirmatory, quantitative research. Furthermore, UGC can be harvested at different points in time and yield panel-type data

to track change and development (Bowen, 2009; Tirunillai & Tellis, 2014). For example, Tirunillai and Tellis (2014) analyzed product reviews to track the importance of different product dimensions over several years. Moreover, and strongly relevant for the present dissertation, the online presence of certain topics suggests how salient they are on the minds of users. The mere fact that beliefs and topics are expressed—without being solicited by a researcher—implies salience (Ksiazek, 2018). Text data is unobtrusive as it has been created without researcher intervention; it is non-reactive as it is not affected by a research process and common method biases (Bowen, 2009; Podsakoff et al., 2003). It avoids social desirability biases due to the research context (Golder & Macy, 2014), an issue that is particularly present in the context of behavior with an ethical dimension, such as sustainable consumption and organic food consumption in particular (Cerri et al., 2019; Hjelmar, 2011), when respondents try to meet the expectations of others. Acquiescence bias can be an issue in surveys and interviews because individuals tend to agree with others (Dillman et al., 2014). However, at the same time, user-generated content can be influenced by impression management toward other users, that is, users consciously or subconsciously seek to influence how they are perceived by others (Hogan, 2010).

UGC is a data source that is abundantly available at little cost and that can encompass a large number of individuals. Thus, insights from text are scalable to the entire population of an online source selected for a study. The potential of UGC, as previously outlined, led to the development of different methodological approaches starting from qualitative content analysis (Kassarjian, 1977; Krippendorff, 2019) to larger data sets analyzed with different computer-aided methods from the field of Natural Language Programming (NLP), which combines artificial intelligence and computational linguistics (Berger et al., 2020). In the marketing and consumer research literature, such computational text analysis methods are denoted as text mining (Netzer et al., 2012), automated text analysis (Humphreys et al., 2018), or computer-aided analysis (Pollach, 2012). Hereafter, the term text mining will subsume any type of UGC analysis.

UGC analysis provides insights into consumer thinking for marketing and consumer research (Balducci & Marinova, 2018; Netzer et al., 2012). Therefore, researchers increasingly seek to tap its potential. Pioneering UGC research in the field of consumer research and digital anthropology, Kozinets (2002) coined the approach of netnography, that is, the systematic analysis of online communities, including discussion content as

well as discussion structure and culture. Ever since, the body of research employing UGC is steadily increasing. Text mining provides a means for researchers and firms to listen to what consumers think about products, product categories, and brands; they can surveil brand positioning; they can learn about which product attributes matter to consumers, and which product-related (societal) issues are on consumers' agendas (Netzer et al., 2012; Tirunillai & Tellis, 2014). In light of this, UGC lends itself to inquire what is on consumers' minds, i.e. what is naturally salient. UGC reveals not only what is discussed online, but also how frequently it is discussed, i.e., the degree of salience, how present a concept is on consumers' minds. UGC brings researchers closer to answering not only the question of *what* consumers do, but *why* they do it (Berger et al., 2020).

1.4 Research aim

This dissertation aims at exploring consumer perception and behavior regarding organic food in the online sphere in the German and US contexts, the largest organic food markets globally. It contributes to consumer research by applying an interdisciplinary, mixed-method research approach that triangulates different data and methods. The thesis uses both qualitative and quantitative data (text and survey data); it integrates methods from consumer and computer science research; and it extends the consumer research perspective by communication research theory.

In particular, this research investigates which beliefs and topics¹ are mentioned by online commenters and their degree of salience. Salient beliefs and topics are accessible associations, product attributes and knowledge that have been stored in memory. It is scrutinized whether the online salience of topics in reader comments is influenced by the news media as agenda setter or vice versa. Furthermore, potential implications from insights on online salience on consumer behavior are explored. To this end, it is analyzed whether, when used as primes, the degree of salience of topics discussed online impacts consumer judgment and decision-making regarding organic food.

¹ The terms „topics“ and „themes“ are used interchangeably in this thesis.

The insights on online salience of organic food beliefs and topics, agenda setting, and behavioral influence are relevant for researchers, policy makers, public opinion polling, as well as practitioners in media and marketing communication.

The present thesis addresses the following research questions:

1. *Which beliefs and topics about organic food are salient to online commenters?*
2. *How can small and large-scale text-mining methods be combined to yield validation?*
3. *Are there agenda setting effects between the media and their audience regarding salient organic food topics?*
4. *Is online salience of topics relevant to consumer behavior?*

Thus, this research focuses on UGC as a source of insight regarding what consumers think and do with respect to organic food. It centers on the following four research gaps and areas.

First, the dissertation examines what consumer say online about organic food to give insights into consumer thinking (Balducci & Marinova, 2018; Netzer et al., 2012). Online commenting can serve as an indicator of which beliefs and topics are naturally salient to consumers. An up-to-date picture of what is on consumers' minds regarding organic food is lacking and the potential of UGC for understanding consumer perception of organic food and related behavior has not yet been sufficiently tapped (Olson, 2017). In addition to taking stock of *what* is salient (which beliefs and topics), this research also accounts for the relative salience of beliefs and topics by counting their frequency of mention.

Second, it is crucial to learn how such text mining can be conducted. There are many novel text analysis techniques yet to be exploited by consumer research (Berger et al., 2020). By triangulating two analysis methods, the established method of content analysis and NLP-based analysis, the dissertation investigates their compatibility and validates the findings of the latter for further applications in this thesis (Renz et al., 2018; Thurmond, 2001; Yu et al., 2011).

Third, Agenda-Setting Theory postulates that the news media influence people's opinions, hence, the public agenda (McCombs & Shaw, 1972). Related studies mostly

focused on political opinion, agenda-setting research regarding consumption issues is scarce. For the context of organic food consumption, Thøgersen (2006) found first agenda-setting evidence; in his study in the Danish context, the valence of media articles was related to self-reported consumer attitudes. This is the first study to investigate agenda setting between news media (i.e., news articles) and their immediate audience (i.e., reader comments) in the field of organic food and sustainable consumption.

Fourth, not only is there a knowledge gap regarding *what* is salient to consumers in online environments, but also *how* online salience of topics relates to consumer judgment and decision-making. Knowledge is stored in associative networks (Collins & Loftus, 1975). Spreading-Activation Theory postulates that upon the presentation of the attitude object or its attributes, individuals automatically activate their attitude toward an object/product, which affects subsequent respective evaluations (Fazio, 1986; Shavitt & Fazio, 1991). This dissertation tests whether priming based on topics salient in the online discussion on organic food (i.e., favorable associations to organic food) can effectively promote pro-organic consumer judgments.

1.5 Structure of the dissertation

This dissertation addresses the aforementioned research questions in four consecutive essays. To provide guidance through the subsequent chapters, Table 1 summarizes the research topic, research question, methodology, and the key findings and contributions of each essay.

The remainder of the thesis is structured as follows: Section 2 will provide the context and the theoretical and methodological background of the dissertation. Subsequently, the four essays composing this thesis will be presented in Section 3. Finally, Section 4 will summarize the key findings and contributions and discuss research, managerial, and policy implications.

Table 1 Overview of essays and their research questions, methodology, key findings, contributions, and triangulation.

	<i>Essay I</i>	<i>Essay II</i>	<i>Essay III</i>	<i>Essay IV</i>
Topic	Consumer beliefs and topics	Validation of two text-mining techniques	Agenda setting	Priming based on online salience of topics
Research Question	<ul style="list-style-type: none"> Which beliefs and topics about organic food are salient to online commenters in German-speaking countries and the US? 	<ul style="list-style-type: none"> How can qualitative content analysis and machine-learning-based analysis methods complement each other? 	<ul style="list-style-type: none"> Are there agenda-setting effects between the news media and their audience regarding salient organic food topics? 	<ul style="list-style-type: none"> Can topics salient online be used to prime consumer judgment and decision-making regarding organic food?
Methodology	<ul style="list-style-type: none"> Content analysis of online comments 	<ul style="list-style-type: none"> Method triangulation of content analysis (<i>Essay I</i>) and filtering approach based on semantic similarity using neural networks 	<ul style="list-style-type: none"> Neural-network-based topic modeling of news articles and reader comments Agenda-setting metrics 	<ul style="list-style-type: none"> LDA topic modeling of online comments Online survey with priming intervention and choice experiment
Key findings and contributions	<ul style="list-style-type: none"> Inventory of organic food beliefs, including several beliefs neglected in previous studies 65 organic food beliefs regarding 22 topics and 4 main topics Online salience differs between beliefs/topics and countries Contributes belief items for survey research and a category system for future text analyses 	<ul style="list-style-type: none"> Identification of 3 evaluation criteria for the selection of appropriate strictness of semantic similarity filtering of sentences Converging results between the two methods Filtering on topic level is superior to filtering on belief level Validity suggests application of filtering to larger unseen organic data set and multiple languages 	<ul style="list-style-type: none"> The news media are agenda-setters for public opinion in the US and Germany Media topic salience influences topic salience in public agenda, but not vice versa Specific events affect agenda diversity and the distance between media and public discussion First study investigating agenda setting w.r.t. organic food topics 	<ul style="list-style-type: none"> Primes based on topics with high online salience increase pro-organic judgment and decision-making. Priming effectiveness depends on product, prime strength, and moderation by attitude strength and buying experience w.r.t. organic food Online salience to be considered for effective marketing and policy communication
Triangulation (within and across essays)	<p>Investigator triangulation</p> <ul style="list-style-type: none"> Cohen's Kappa confirmed intercoder reliability, i.e., the application of the category system among the two researchers categorizing the text was consistent 	<p>Method triangulation</p> <ul style="list-style-type: none"> Validation of the semantic similarity filtering by comparing results to the manual categorization in <i>Essay I</i> (convergent validity) The use of neural networks for semantic similarity was validated for application in topic modeling in <i>Essay III</i> 	<p>Data and method triangulation</p> <ul style="list-style-type: none"> Identified 8 topics match conceptually with topics from <i>Essay I</i> (convergent validity) Topic-correlation, agenda distance, and agenda diversity metrics suggest close relationship between the topics discussed in news articles and reader comments (convergent validity) 	<p>Data and method triangulation</p> <ul style="list-style-type: none"> Identified 18 topics match conceptually with those from <i>Essays I</i> and <i>III</i> (convergent validity) Behavioral relevance of online salience is demonstrated (external validity)

Source: Own table.

2 Background

This section presents the empirical, theoretical, and methodological background of the dissertation. Section 2.1 will introduce the four perspectives on user-generated content taken in this dissertation. Section 2.2 will proceed with an overview of the organic food market and its consumers. Section 2.3 will present the research approach of this thesis. Data collection and analysis will be described including a review of text-mining methods as well as of measures of consumer perception and choices.

2.1 The four perspectives of the dissertation

This essay-based dissertation takes four perspectives on UGC in the context of organic food. These perspectives are the following and will be described in the subsequent subsections: An exploratory, qualitative perspective to create an inventory of organic food beliefs voiced in online comments (*Essay I*), a methodological perspective on two techniques to analyze textual UGC (*Essay II*), an agenda-setting perspective on topic influences between the news media and their audience (*Essay III*), and a perspective on the behavioral relevance of online salience of topics (*Essay IV*).

2.1.1 Organic food beliefs in online comments – A qualitative perspective (Essay I)

The organic food market has been growing continuously in the past decades, which has led to a large body of research on organic food consumption (for literature reviews see, e.g., Hughner et al., 2007; Hemmerling et al., 2015; Rana & Paul, 2017). To spur organic food consumption, it is key to understand consumers' perception of organic food and their buying motives. Most of current and recent studies are survey research eliciting participants' agreement to a limited number of typical beliefs about organic food (e.g., Lee & Hwang, 2016; Nuttavuthisit & Thøgersen, 2017). In recent years, only little qualitative research has been conducted to explicitly and comprehensively confirm and update the beliefs currently held by consumers in Europe or the US (Aschemann-Witzel

& Niebuhr Aagaard, 2014; Vega-Zamora et al., 2014). The qualitative research that has been carried out often focuses on specific aspects of organic food such as healthiness or origin (Ditlevsen et al., 2020; Pedersen et al., 2018). Furthermore, researchers are only starting to exploit the potential of UGC to investigate consumer perception of organic food and related behavior. First examples are an analysis of organic food beliefs of US news readers (Olson, 2017) and a keyword and a semantic network study investigating Tweets on organic food from Mexico and South Korea (Meza & Park, 2016). *Essay I* seeks to further tap the potential of UGC and meet the lack of up-to-date exploratory research into consumer beliefs in major organic food markets. It pursues the following research question: *Which beliefs and topics about organic food are salient to online commenters in German-speaking countries and the US?* The essay delivers a detailed inventory of beliefs, a categorization of beliefs into broader discussion topics, and insights on how the salience of beliefs and topics differs between the two regions.

2.1.2 Organic food beliefs in online comments – Combining text-mining techniques (*Essay II*)

As UGC data becomes more ubiquitous, consumer researchers have continuously developed and adopted techniques to analyze such textual online data to yield consumer insight (Balducci & Marinova, 2018; Vidal et al., 2018). However, different text-mining techniques come with different limitations. Whereas classical content analysis is performed manually and delivers high analytical depth (Krippendorff, 2019), it is bounded by the amount of data that can be analyzed. In contrast, automated techniques based on pre-defined dictionaries and word-count (referred to as “automated text analysis”) can handle large amounts of data, but are limited in terms of analysis depth as they rely on the frequency of single words (Berger et al., 2020). Going beyond such automated text analysis, the computer science field of NLP (a combination of artificial intelligence and computational linguistics) has developed both supervised and unsupervised machine-learning approaches, which are slowly adopted by consumer researchers (Berger et al., 2020; Vidal et al., 2018). On the one hand, supervised machine learning requires large amounts of labeled data (i.e., “coded data” in qualitative research terms) to train algorithms that classify text into pre-defined categories. Such an approach was trialed in an early stage of this dissertation project. Several researchers labeled a data set of approximately 10,000 sentences with a category system based on *Essay I*.

Nevertheless, the amount of labeled data proved insufficient to yield reliable classification results using supervised machine learning. Unsupervised machine learning for text analysis, on the other hand, comprises topic modeling, such as Latent Dirichlet Allocation (LDA, Blei et al., 2003). Such probabilistic models of co-occurrence of words (see preliminary text-mining study in *Essay IV*) are well-established and validated, however, a limitation lies in the ignorance of semantic context. In contrast, recent advances in neural network and deep learning research allow to take semantic context into account. Algorithms have been pre-trained on large web-scale corpora and encode the semantic context in word or sentence vectors/embeddings (Cer et al., 2018; Mikolov et al.). Examples are the Universal Sentence Encoder (used in *Essay II*), which is the basis for Google's translation services, and XLING (used in *Essay III*). This dissertation presents two possible applications of such pre-trained neural networks in consumer research. While *Essay III* uses sentence embeddings for topic modeling, *Essay II* employs sentence embeddings to filter sentences based on semantic similarity. It seeks to replicate the results of *Essay I*, and thus triangulates two text-mining techniques. Previous research has discussed and argued for the compatibility of qualitative and machine-learning-based text-analysis techniques (Isoaho et al., 2019; Yu et al., 2011). Many studies rely on NLP techniques to extract the data from the web and, if applicable, compose smaller purposive samples (Shahin, 2016). In addition to mere data collection, computational techniques are incorporated in data analysis. For instance, Renz et al. (2018) have combined content analysis and word count analysis using the LIWC dictionary. *Essay II* pursues a methodological research question by empirically investigating the compatibility: *How can qualitative content analysis and machine-learning-based analysis methods be combined to yield validation?* *Essay II* develops a process regarding how the analysis depth of a manual content analysis can be coupled with the scope of pre-trained sentence embeddings, and suggests criteria to evaluate the filtering of sentences based on semantic similarity.

2.1.3 Agenda setting in organic food (*Essay III*)

The media play a determinant role in public opinion formation. This relationship is explained by the Agenda-Setting Theory (McCombs & Shaw, 1972). The influence between the media and public agendas has been frequently tested by comparing topics and sentiment of news articles of selected media outlets to those stated by their audience

in reader comments sections, on different social media platforms, or in attitudinal surveys (Abdi-Herrle, 2018; Conway-Silva et al., 2018; Gerber et al., 2009). Most of this research has focused on investigating political opinion. Research on agenda setting in a consumption context is scarce. Thøgersen (2006) has provided the first and to date only evidence for the agenda-setting mechanism in the context of organic food consumption by indicating a relationship between news article sentiment and opinion polls. The media could play a crucial role in shaping more sustainable consumption patterns (Trivedi et al., 2018). *Essay III* seeks to shed light on the media as a potential predictor of organic food topics in consumer opinion. The research question of *Essay III* reads: *Are there agenda-setting effects between the media and their audience regarding salient organic food topics? Essay III* analyzes the relationship between discussion topics in news articles and reader comments for two major German and US newspapers. Further, the semantic distance and diversity between media and public agendas is compared.

2.1.4 Priming organic food behavior based on online salience (*Essay IV*)

With the increasing exploitation of UGC for consumer insight comes the question of how online consumer statements relate to consumer behavior. The sparse previous research in this regard has focused on the predictive power of product reviews on sales, and on measuring posting volume (for a meta-analysis, see Babić Rosario et al., 2016) but not extracting topical content. For example, in the context of organic food, it was found that the number of reviews predicts product sales on the Chinese e-commerce platform *Taobao* (Lyu & Choi, 2020). In contrast, UGC such as online comments from discussion boards or news websites have been little studied in terms of the relationship between topics discussed on these sources and consumer evaluation and behavior. Such UGC delivers detailed insights on which topics and product attributes are salient on consumers' minds and how the degree of salience of topics differs. It is thus an indicator for the natural salience of topics and attributes associated with a product (Ksiazek, 2018). The underlying theoretical conceptualization is that consumers store product knowledge in associative networks (Collins & Loftus, 1975) and can activate the attitude toward a product when primed with the product itself or an associated attribute (Fazio et al., 1986; Shavitt & Fazio, 1991). Against the background of this Spreading-Activation Theory, *Essay IV* seeks to answer the following research question: *Can topics salient online be used to prime consumer judgment and decision-making regarding organic food?* In *Essay*

IV, topics of high and low online salience are first identified via text mining. Subsequently, an online priming experiment tests whether the degree of online salience matters for three outcome variables: choosing organic alternatives in two choice experiments and the acceptance of policies supporting organic farming. *Essay IV* analyzes the salience of topics related to organic food in online environments can be used to promote pro-organic behavior. For example, such online salience nudges could for effective communication by marketing practitioners or policy makers.

2.2 Organic food markets and organic food consumption

Organic food markets and labels. Worldwide, the organic food sector experienced significant growth in the last decades, propelled by the implementation of national and supra-national organic production standards (Sahota, 2018). The present research focuses on the German and US contexts as the most important organic food markets globally (Willer et al., 2020).

In Europe, the organic movement began in the 1920s with the establishment of the first organic farming associations that still exist today and issue private organic labels, such as Demeter and Bioland (Organic Food Production Alliance [BÖLW], n.d.). In 1992, EU-wide regulation for organic farming and labeling was adopted. In addition to public organic labels issued at the national level (e.g., the German organic label introduced in 2001), the EU organic label was introduced in 2010 (European Commission, 2020a). EU regulation defines organic farming as follows:

“Organic farming is an agricultural method that aims to produce food using natural substances and processes. This means that organic farming tends to have a limited environmental impact as it encourages: the responsible use of energy and natural resources; the maintenance of biodiversity; preservation of regional ecological balances; enhancement of soil fertility; maintenance of water quality. Additionally, organic farming rules encourage a high standard of animal welfare and require farmers to meet the specific behavioural needs of animals.” (European Commission, 2020a)

In the US, organic farming has become increasingly well-known from the 1960s onwards (Pearson et al., 2010). In 1990, production standards for organic food were defined by the Organic Food Production Act, but it was not until 2002 that public organic certification was introduced under the National Organic Program (USDA Economic Research Service, 2020). Unlike the EU, the USDA (United States Department of Agriculture) organic label distinguishes four certification categories: ‘100 percent organic’; ‘organic’ (minimum 95% of organic ingredients); ‘made with organic ingredients’ (minimum 70% of organic ingredients); and ‘specific organic ingredients’ (less than 70% organic ingredients). While only the first two categories carry the USDA organic seal, all four may use the term ‘organic’ (USDA Agricultural Marketing Service, 2016; USDA Economic Research Service, 2020).

Today, the US is the most important organic food market globally accounting for 42% of global organic food sales, followed by Germany with 11% (Willer et al., 2020). As shown in Figure 1, organic retail sales have been continuously increasing in the past years. In 2018, the relative market share of organic food of total food sales was 5.3% in Germany and 5.7% in the US (FiBL, 2020, see Figure 1). In the same year, organic retail sales have grown by 5.5% in Germany and by 5.9% in the US (FiBL, 2020).

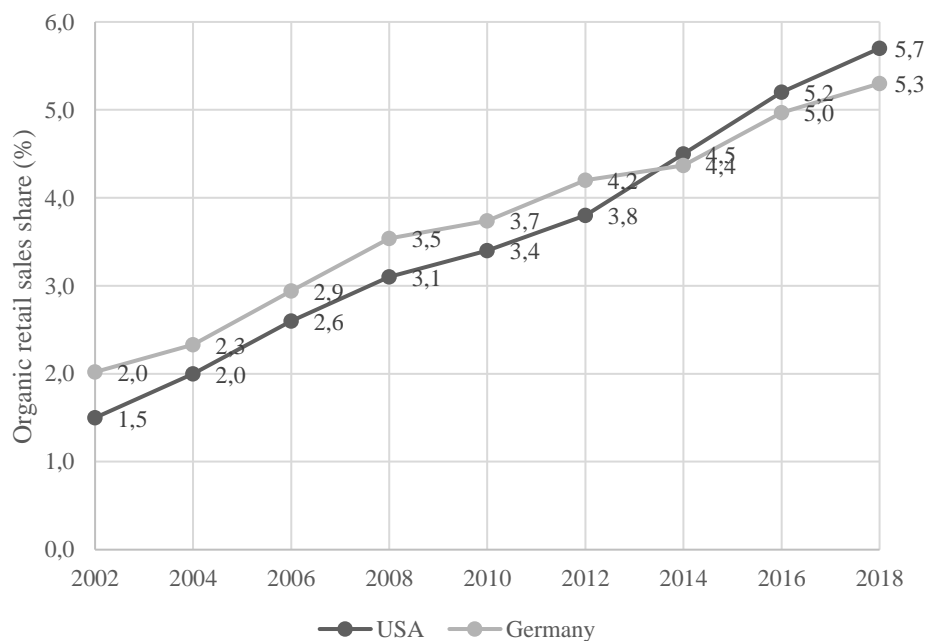


Figure 1 Organic retail sales share (%) from 2002 to 2018 in Germany and the USA. Source: Own illustration based on FiBL (2020).

Organic food consumption. Alongside with growing organic demand, a large body of research has emerged, which investigates consumer perception of organic food and the antecedents of organic food purchases. Frequently investigated determinants of organic food purchases are beliefs, attitude, intention, and socio-demographics (Chen, 2020; Hempel & Hamm, 2016; Janssen, 2018; Klöckner, 2013). Consumer beliefs as the drivers of positive attitudes, buying intention, and thus preference for organic food have been frequently investigated (for reviews see Aertsens et al., 2009; Hemmerling et al., 2015; Hughner et al., 2007; Kushwah et al., 2019; Pearson et al., 2010; Rana & Paul, 2017; Schleenbecker & Hamm, 2013; Yiridoe et al., 2005). These reviews summarize a wide range of beliefs about different organic food attributes and associated topics. Beliefs are related to the product itself (e.g., healthiness, food safety, price, taste, quality and appearance, nutritional value, naturalness, genetically-modified organism free, availability, local origin), organic labeling and certification, production (e.g., environmental impact, animal welfare), and the societal role of organic farming (e.g., food security). Evidence on socio-demographic factors driving organic food purchases is mixed. While some studies have found no significant relationship, others have shown higher levels of income and education, children in the household, and lower age levels to positively influence the level of organic food purchases (for an overview, see Janssen, 2018). Moreover, there are further determinants of organic food behavior in particular and environmental behavior in general stemming from different consumer behavior theories (see Section 2.3.3.1), such as values, personal and social norms, perceived behavioral control, and habits (Chen, 2020; Janssen, 2018; Klöckner, 2013).

However, it has to be noted that positive attitudes toward organic food, buying intention, and self-reported behavior do not always translate into or correspond to actual organic food purchases. In a consumer survey, 32.2% of Germans (30.3% of US citizens) indicated that more than 50% of their food purchases was labeled organic, with only 6.6% (14.9% in the US) stating they never buy organic food (Kitchen Stories, 2020a, 2020b). However, the actual yearly per-capita consumption of organic food was only €132 in Germany and €125 in the US in 2018 (FiBL, 2020). Using survey and household panel data for Germany from 2008, Janssen (2018) showed that while 20% of the households had positive attitudes toward organic food, only a very small share of all households (3%) actually spent more than 20% of their food budget on this product category. This attitude-behavior gap has been frequently documented and investigated (Chekima et al.,

2017; Padel & Foster, 2005; Schäufele & Hamm, 2018; Vermeir & Verbeke, 2006). Methodological reasons for such a gap could be biases rooted in social desirability or acquiescence (see Section 1.3). Moreover, the discrepancy is explained with both supply-side factors, such as price and availability (Aschemann-Witzel & Niebuhr Aagaard, 2014), as well as demand-side factors, such as subjective norms and values (Thøgersen et al., 2017). Furthermore, consumers tend to overestimate the actual share of their organic purchases due to confusion about what organic products actually are, for example by taking products for organic which are not (Harper & Makatouni, 2002; Hughner et al., 2007). It is thus crucial to better comprehend consumer perception and understanding of organic food. Beliefs and attitudes are meaningful to understand organic food behavior: it has been shown that—in spite of an existent gap—attitudes and actual purchases are driven by the same determinants (e.g., perception of healthiness, environmental protection, quality, and price) (Janssen, 2018).

2.3 Research approach

This dissertation aims to understand the online consumer perception of organic food, influential factors such as the media, and their relevance for consumer judgment and decision-making. It pursues an interdisciplinary, mixed-methods approach to reach this aim and answer the research questions posed in Section 1.4. The thesis collects textual UGC from online comments from different social media platforms as well as online news article data to analyze salient beliefs and discussion topics. Furthermore, survey data are employed to measure consumer attitude, voting behavior, and choices.

2.3.1 Overview of data collection and analysis

The data used in this dissertations stem from a large data set of UGC on organic food (hereafter referred to as *Text Data Corpus*) as well as two consumer surveys (hereafter referred to as *Preliminary Survey* and *Main Survey*).

The *Text Data Corpus*. The *Text Data Corpus* was built in collaboration with the *Research Group Social Computing* of Technical University of Munich. Web crawling was employed to collect publicly available user comments from online discussion boards (e.g., *quora.com*, *reddit.com*) as well as reader comments and news articles from online

news platforms (e.g., *nytimes.com*, *spiegel.de*) from the US and German-speaking countries, that is, Germany, Austria, and Switzerland (for an overview of sources in the *Text Data Corpus* see Table 2 in Section 3.1.3.1). The selected sources provide several advantages over social network data, such as Facebook or Twitter, which suffer from data privacy and accessibility issues, impression management due to non-anonymous profiles, and, in the case of Twitter, limited comment length (Markham et al., 2018; Steinert-Threkeld, 2018; Ziegele, 2016). For general characteristics of textual UGC see Section 1.3. The *Text Data Corpus* was composed in November 2017 and updated in February 2020; it considered all user comments and news articles (if applicable) available on the respective online sources since January 2007. After data acquisition, the following data preprocessing steps were performed on the *Text Data Corpus*: First, the data were tokenized, that is, the text was broken into smaller units of analysis such as words and sentences. Second, the data were cleaned from meaningless text (e.g., URLs). Third, irrelevant articles and comments were filtered out by building a Naïve Bayes Classifier, which was trained on a labeled data set and cross-validated. Fourth, common spelling mistakes were eliminated. Fifth, for some UGC analyses, stop words (e.g., common words such as “and”, “or”, “a”) were removed and words were lemmatized (i.e., words were reduced into their dictionary form by removing inflectional forms).

The essays in this dissertation draw on sub-samples of the previously described *Text Data Corpus*. *Essay I* and *II* rely on the same random sub-sample of online comments from the *Text Data Corpus*. Thus, they consider a wide range of sources from several discussion boards and news websites from the US (N=1069) and German-speaking countries (N=1094) published between January 2007 and November 2017. Data in *Essay I* are analyzed via qualitative content analysis. *Essay II* combines this content analysis with a machine-learning-based topic modeling approach using neural networks for sentence embeddings. *Essay III* compares the reader comments and news articles of two major German and US news websites. It uses all reader comments and news articles from *spiegel.de* (comments: N=63,379; articles: N=568) and *nytimes.com* (comments: N=41,320; articles: N=534) available from January 2007 to February 2020. Data are analyzed using a topic modeling approach that clusters sentence embeddings. In a preliminary text-mining study, *Essay IV* makes use of a sub-sample of the *Text Data Corpus* to inform an online priming experiment conducted in the German context. The preliminary text-mining study considers all reader comments from the German news

website *spiegel.de* available from January 2007 to December 2017 (N=63,379). Topics are modeled with a probabilistic topic model using the LDA algorithm.

Survey data. In addition to text data, *Essay IV* also used survey data. A *Preliminary Survey* was conducted among German consumers in January and February 2019. Using convenience sampling, participants (N=224) were recruited via social networks and *SurveyCircle*. The *Preliminary Survey* contributed descriptive measures of Likert-type self-reported importance of different organic food topics to *Essay IV*. The *Main Survey* for *Essay IV* was carried out in October 2019 in cooperation with the market-research company *bilendi*. The sample (N=1118) contained quotas representative of the German adult population in terms of gender, age, schooling, professional education, and income. The *Main Survey* comprised a priming experiment and collected choice data. Choice data are analyzed with a multinomial logit model using latent class analysis, where the priming conditions and the control group were imposed as latent classes. Additionally, the *Main Survey* elicited Likert-type measures on attitude and policy acceptance.

2.3.2 Text mining

In the past decades, a multitude of text-mining methods² has been developed by different research disciplines, such as the social sciences, linguistics, and computer science. With the ever-increasing availability of large amounts of text data, text researchers have sought for more automation in text analysis, and—so has this dissertation—have integrated advances from computer science, more precisely, from natural language processing (for an overview of methods see Balducci & Marinova, 2018; Berger et al., 2020; Hartmann et al., 2019; Vidal et al., 2018).

Overview of text-mining methods. Text analysis methods can broadly be categorized as follows. First, content analysis is a long-standing qualitative research method in the social sciences including consumer research (Kassarjian, 1977; Krippendorff, 2019). Meaning is extracted from text by one or several researchers manually categorizing text. Content analysis can comprise the mere identification of categories or may be quantified by counting the occurrence of categories. Content analysis allows for in-depth extraction of

² The terms text mining and text analysis are used interchangeably.

meaning from text, however, it is usually limited to small data samples due to required time and human resources.

Second, in computer science, the field of NLP is dedicated to the analysis of language spoken or written by humans (Manning & Schütze, 2005). NLP is a sub-division of artificial intelligence and computational linguistics. Social scientists have adopted NLP methods of automated text analysis that are based on word count using dictionaries (Berger et al., 2020). Different dictionaries have been composed, for example, LIWC (Tausczik & Pennebaker, 2010). For a text to be analyzed, the respective software automatically counts the occurrence of the words contained in the dictionary. Whereas such word-count-based methods allow for the analysis of large samples, the depth of analysis is limited as only words encoded in the dictionary are considered and the semantic context of words is disregarded.

Third, NLP has been revolutionized by machine-learning algorithms, which largely replaced dictionary and grammar-based approaches. LDA (Blei et al., 2003) is one of the most frequently employed algorithms to model the semantics, that is, topics, in a text. It is a bag-of-words technique, where a probabilistic model of words and their co-occurrence is created disregarding grammar or order of words. Compared to dictionary approaches, LDA topic modeling has been shown to cover content in a more detailed manner as it is not reliant on pre-defined dictionaries (Guo et al., 2016). However, content analysis, dictionary-approaches, and topic modeling are ultimately count measures. Latest advances using neural networks have introduced semantic similarity measures. Neural networks allow for a more precise consideration of the semantic context, in which words or sentences occur. These techniques use word embeddings (e.g., word2vec; Mikolov et al.) or sentence embeddings (e.g., Universal Sentence Encoder; Cer et al., 2018). Similarity between word or sentence embeddings is commonly measured with cosine similarity (used in *Essays II* and *III*) or the Jaccard index (Berger et al., 2020). Neural network approaches using sentence embeddings suggest themselves as particularly insightful for consumer research: They conserve the analytical detail of qualitative research (*Essay I*) as they work on sentence level and thus consider original consumer phrasings as opposed to word count or bag-of-words measures such as LDA, which is one of the most frequently employed NLP methods in consumer research (Büschken & Allenby, 2016; Timoshenko & Hauser, 2019). Retaining as much context as possible

delivers more insight into consumers' concerns and needs and the required product attributes to satisfy them (Timoshenko & Hauser, 2019). Moreover, they can be applied in a multilingual manner as demonstrated for German and English in *Essay III*. The performance of neural network approaches is assumed to continuously improve. Further and more off-the-shelf NLP solutions for application in consumer research can be expected.

Text-mining in this dissertation. The present work employs and combines several of the outlined text analysis methods, their application will be briefly described, motivated, and compared in the following.

Different text-mining techniques were employed in this dissertation to seek validation between methods and to adapt to the requirements of the research question of each essay. *Essay I* sought for an in-depth analysis yielding a fine-grained category system of organic food beliefs and topics that also considers beliefs and topics delivered by previous literature. For this purpose, manual, qualitative content analysis was the text-mining technique of choice. It also delivered a baseline to compare subsequent automated analyses against. *Essay II* tested whether an automated text analysis can replicate the results of *Essay I*. Only a neural network approach using sentence embeddings for semantic similarity could be expected to reproduce the level of detail provided by the category system of the manual content analysis as other techniques work on word level. *Essay III* aimed at analyzing topical relationships between news articles and reader comments. *Essay II* had proven the usefulness of sentence embeddings to filter text on organic food. Given this validation of using sentence embeddings, *Essay III* chose a clustering of sentence embeddings as a promising topic modeling technique. The clustering of sentence embeddings considers the semantic context of words and was thus found to deliver a more accurate representation of topics than standard LDA topic modeling algorithms. In contrast, *Essay IV* solely required a preliminary study to inform the subsequent priming experiment about topics of high and low salience. To this end, the computationally less intense and widely validated LDA topic modeling was deemed suitable.

The foundation of any text analysis is the classification of words or text segments (i.e., the units of analysis) into classes or categories (Krippendorff, 2019)—here beliefs and topics. Different units of analysis are employed in the studies of this dissertation: single

words and their co-occurrence with other words (*Essay IV*), sentences (*Essay I-III*), and short paragraphs (*Essay I*). Therefore, other than in traditional survey research, where beliefs or topic importance are elicited as the agreement to given survey items, *Essays I* and *II* measure beliefs by categorizing verbal consumer expressions based on their semantics into different belief categories (e.g., “organic farming protects the environment”) and topic categories (e.g., “environment”). These categories both emerged from the data as well as were derived from extant literature. In *Essay I*, this categorization is performed manually in a content analysis. Here, text was categorized by two coders. To ensure consistent application of the categories, intercoder reliability is calculated between the two coders using Cohen’s Kappa (Lombard et al., 2002). *Essay II* seeks to replicate the categorization results of the content analysis of *Essay I* by filtering similar sentences based on their semantic similarity. Using the neural network algorithm of the Universal Sentence Encoder (Cer et al., 2018), the semantic similarity of a sentence is encoded by transforming each sentence into a multidimensional sentence vector (also referred to as sentence embedding). The semantic similarity between two vectors can be compared by measuring their angular distance. Here, the cosine similarity measure was used. To filter the sentences into different categories, each belief category is compared to each sentence in the data set and a threshold for similarity is applied. In *Essays III* and *IV*, text was categorized into topics, i.e. the reference to a specific topic (e.g., “environment”), leaving aside the valence or evaluation of this topic as it is commonly contained in a belief. In NLP, the categorization of text into topics is referred to as topic modeling. To model the topics, *Essay III* used semantic embeddings (like *Essay II*), which were clustered based on semantic similarity using k-means clustering. As sentence embeddings consider the context of not only words but also sentences, it is a more elaborate technique than the topic modeling applied in *Essay IV*. In *Essay IV*, the widely-used LDA algorithm was applied.

The different text-mining studies—and in particular the content analysis in *Essay I*—comprise both the identification of beliefs and topics (*what* is salient) and the quantification of the occurrence of these beliefs and topics (*how* salient). Counting gives an indication for relative importance of concepts. However, little mentioned beliefs and topics are nevertheless meaningful to understand consumer thinking. The chosen unit of measurement (to count occurrence of the respective units of analysis) depended on the research aim of the essays. In *Essays I* and *II*, the unit of measurement was the simple

count of the frequencies of units of analysis per belief and topic category. The unit of measurement in *Essay III* was the relative topic distribution, that is, the relative weight of a topic compared to the other topics represented on a media or public agenda. Finally, *Essay IV* used the share of comments making mention of a given topic as unit of measurement.

Metrics applied to text-mining results in this dissertation. Different metrics were applied to further analyze the text-mining results. In *Essay I*, global chi-square was employed to detect significant differences in category frequencies between the German and US contexts. Correlation analyses were performed in *Essay II* to compare the similarity of results of the content analysis and the neural-network technique as well as in *Essay III* to analyze the relationship between topic distributions in articles and comments with and without time lag. Furthermore, *Essay III* showed the fruitful integration of metrics from information and probability theory into agenda-setting research, in particular, Shannon Entropy for agenda diversity and Jensen-Shannon Divergence for agenda distance.

2.3.3 Theories and measures of consumer perception and behavior

Using UGC for consumer insight is a rather novel approach. By applying different text-mining methods for UGC to detect salient beliefs and topics, *Essays I-IV* take an exploratory angle to investigating consumer perception, in particular consumer beliefs and associated topics salient to consumers. By contrast, consumer perception and behavior are long-standing subjects of behavioral theories and related measures from marketing and consumer research as well as consumer economics (Peter & Olson, 2010; Phillips et al., 2002). In addition to a preliminary text-mining study, *Essay IV* employs a confirmatory research approach and—by measuring attitude, voting behavior, and consumer choices—combines both the perspectives of marketing and consumer research and consumer economics. The following sub-sections will give an overview of the theoretical background and measurement of consumer beliefs, attitudes, and choices.

2.3.3.1 Beliefs and attitudes

Theories. The measurement of salience of beliefs and topics related to organic food associated in consumers' memory is the common denominator of all four essays.

Consumers store acquired product knowledge, that is beliefs related to an object, product, or concept in memory. These beliefs form associative networks of linking all meanings related to the product (Bettman, 1979; Collins & Loftus, 1975). Beliefs are defined as knowledge and meanings associated with a product, its production, and its consumption (Fishbein, 1963; Peter & Olson, 2010). Beliefs about an object and attitude toward the object as the sum of all salient and evaluated beliefs are considered early antecedents of consumer behavior by numerous behavioral theories, for example, the Multi-Attribute-Attitude Model (Fishbein, 1963) and the succeeding Theory of Reasoned Action (Ajzen & Fishbein, 1980; Fishbein & Ajzen, 1975) and Theory of Planned Behavior (Ajzen, 1991), the Spreading-Activation Theory (Collins & Loftus, 1975), the Attitude-Activation (Fazio, 1986) and the MODE Model (Fazio, 1990). Further behavioral theories involving beliefs are the Value-Belief-Norm Theory (Stern et al., 1999; Stern, 2000), the Alphabet Theory (Zepeda & Deal, 2009), which integrates the Value-Belief-Norm Theory and the Attitude-Behavior-Context Theory (Guagnano et al., 1995), and the Behavioral Reasoning Theory (Westaby, 2005).

This thesis investigates what people post online; it builds on the assumption that what is mentioned online is a natural indicator of which beliefs and topics are salient to users. First, the dissertation focuses on taking stock of which beliefs and topics are salient to online users (*Essays I-IV*). Against this theoretical background, this thesis defines salient beliefs and topics as all (evaluated) accessible associations and product attributes that easily spring to mind or are prominent when seeing or thinking about a product (Ajzen & Fishbein, 2005; Fazio, 1990; van Ittersum et al., 2007). While related to the notion of perceived importance, salience especially entails accessibility of associations (top-of-mind awareness) (Takeshita, 2006). Fazio's Attitude-Activation Model postulates that attitudes are summary evaluations of an object that stem from past behaviors and experiences (Fazio, 1986). In the sense of spreading activation, these extant attitudes can be automatically activated upon the presentation of the attitude object or its attributes. For example, reading a news article on organic food online activates the associative network on organic food and the respective attitude stored in memory. The stronger the evaluated associations (i.e., beliefs) in memory, the more accessible (i.e., salient) the attitudes and the more likely is their automatic activation (Fazio et al., 1986; Fazio & Olson, 2003). Considering the legacy of the Theory of Planned Behavior, Fazio further extended the Attitude-Activation Model to a dual process model, the MODE Model (Fazio, 1990),

which was also recognized by Ajzen and Fishbein (2005). He argues that—as previously described—attitudes are automatically activated from memory, except if consumers are motivated and have the opportunity to deliberate further, then they will recur to reasoned action as postulated by Theory of Planned Behavior (e.g., when no prior attitude was stored in memory).

Second, the dissertation aims at measuring the overall degree of salience in terms of how frequently beliefs and topics are mentioned in the investigated UGC sources (*Essays I-IV*). Previous survey research has frequently measured accessibility in terms of response latency, for example, using the Implicit Association Test (Maison et al., 2004).

Third, the dissertation investigates whether the relationship between online salience of topics and consumer judgment (*Essay IV*). Salient information is “attention grabbing” and can thus be used to nudge consumers toward a certain behavior (Reisch & Zhao, 2017; Thaler & Sunstein, 2009; Tversky & Kahneman, 1974). Building on the spreading-activation model of associative memory and priming theory (Fazio et al., 1986; Janiszewski & Wyer, 2014; Minton et al., 2017), *Essay IV* hypothesizes that the presentation of a salient topic in the context of organic food activates the associative network of organic food, increases the accessibility of attitude toward organic food, and thus can influence subsequent consumer judgment and decision-making regarding organic food.

Measurement. Confirmatory studies applying the aforementioned theories measure consumer beliefs, attitude, and voting behavior in terms of liking. In Likert-type questions, participants are usually asked to rate their agreement with a number of belief, attitude, or intention items (Ajzen, 2002). *Essay IV* applies such rating scales to assess attitude valence and attitude strength (Pomerantz et al., 1995). Additionally, *Essay IV* employed a rating scale to measure the acceptance of a policy supporting organic agriculture (Alphonse et al., 2014; Tonsor et al., 2009)

Prior to eliciting beliefs in the mentioned confirmatory fashion, exploratory research is needed to identify consumer beliefs (Ajzen, 2002). To do so, conducting focus groups and interviews with consumers is the most common approach. For instance, Padel and Foster (2005) and Harper and Makatouni (2002) used focus groups and laddering interviews to explore consumer beliefs about organic food. Aschemann-Witzel and Niebuhr Aagaard (2014) inquired the thoughts of organic food shoppers at the point of

sale. More seldom, UGC is exploited to investigate beliefs, although it particularly lends itself to deliver an up-to-date picture of current consumer beliefs and also uncover perceptions little considered in previous literature as they might remain unmentioned in researcher-guided interviews or focus groups (see Section 2.3.2). An exception poses Olson (2017), who analyzed organic food beliefs stated in reader comments of US news outlets. While his focus is on detecting the change in certain organic food beliefs, the purpose of *Essay I* is to create a comprehensive inventory of current organic food beliefs voiced in a variety of online platforms. As opposed to beliefs, attitude is rarely elicited in an exploratory fashion through qualitative data analysis. However, text analysis is providing novel sources of insight for the measurement of attitude. The valence or sentiment in UGC textual statements can be used as an indicator of overall attitude (Olson, 2017). The value of sentiment analysis of text is increasingly acknowledged in consumer research (Ordenes et al., 2017). Automated sentiment-analysis tools as an indirect measure of liking using text data are continuously developed (Visalli et al., 2020). Further advances in sentiment analysis in the disciplines of consumer analytics and Natural Language Processing can be expected to influence the future measurement of attitude.

2.3.3.2 Choices

In addition to attitude measures and text-mining, measuring consumer preferences is at the core of *Essay IV*. The concept of consumer preferences is rooted in microeconomic theory and refers to the liking of one product or attribute more than another. Preferences can be elicited via discrete choice analysis, which explains the heterogeneity in the behavioral response of consumers (Louviere et al., 2010). Whereas stated preferences are given in hypothetical situations like a choice experiment (*Essay IV*), revealed preferences can be collected from choices made in real market situations (Louviere et al., 2010). In choice experiments, consumers choose between alternatives. According to the Theory of Random Utility Maximization (McFadden, 1974), consumers know their preferences and choose the alternative offering the best trade-off relative to the other alternatives offered (Manski & McFadden, 1981). Lancaster's (1966) approach to consumer theory contributed that alternatives are characterized by attributes providing utility to consumers. Attributes have different levels (e.g., different levels of price). Choice analysis focuses on the relative utility of alternatives in a given choice set. Going beyond mere utility maximization, choice theory has integrated behavioral and psychological perspectives,

which suggested that preferences are not solely pre-determined, but also result from heuristic rules, decision context, and emotions (Adamowicz et al., 2008; Loewenstein, 2000).

In a choice experiment, participants face several choice scenarios (i.e., choice sets), where they decide between alternatives composed of different attribute levels. Due to the trading-off of attributes, choice experiments are considered a more reliable estimate to evaluate consumer utility compared to attitudinal data (Louviere et al., 2010), in particular for decision contexts with an ethical dimension and prone to biases (Auger & Devinney, 2007). Choice experiments have been frequently applied to study consumer behavior with regard to food and eating, and organic food in particular (e.g., Gracia & Magistris, 2008; Hempel & Hamm, 2016; Thøgersen et al., 2019).

In the following section, the four essays composing the dissertation will be presented.

3 Four essays

3.1 Using online comments to explore consumer beliefs regarding organic food in German-speaking countries and the United States (*Essay I*)³

Abstract

Consumer beliefs play an important role in explaining consumer behavior. This exploratory study aims at building an inventory of consumer beliefs about organic food. To reach this objective, we conducted a content analysis of online comments about organic food posted on news websites and forums in German-speaking countries (n=1,094) and the United States (n=1,069). Such user-generated content has emerged as an abundant source of insight for consumer research, although very little has been exploited in organic food consumption research.

The main result of this study is a comprehensive category system of 65 organic food beliefs and their relative frequencies. The category system reflects the large variety of beliefs and their differing salience within and across the two regions studied. We discuss the relevance of our category system for future survey, experimental, and textual research, as well as for marketing practitioners and policy-makers.

Keywords: Consumer beliefs; organic food; social media; content analysis.

³ *Essay I* has been published as Danner, H.; Menapace, L. (2020): Using online comments to explore consumer beliefs regarding organic food in German-speaking countries and the United States. *Food Quality and Preference*, 83, 103912. DOI: 10.1016/j.foodqual.2020.103912. Adaptations to fit the format of this thesis have been made.

CRedit authorship contribution statement: **Hannah Danner:** Conceptualization, Methodology, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing, Visualization, Project administration, Funding acquisition. **Luisa Menapace:** Conceptualization, Writing – original draft, Writing – review & editing, Supervision, Funding acquisition.

3.1.1 Introduction

The organic food sector has grown considerably around the world over the last 20 years, spurred by the development of national organic standards (Sahota, 2018). As such, there is considerable research interest in understanding the respective consumer behavior (Hemmerling et al., 2015; Hughner et al., 2007). The Theory of Planned Behavior (Ajzen, 1991) proposes that beliefs play an important role in explaining consumer behavior. According to this theory, consumers hold beliefs about a product and evaluate those beliefs to form an attitude toward the product, which in turn influences purchase intention and behavior. In a meta-analytic structural equation model, Scalco et al. (2017) showed the robustness of the Theory of Planned Behavior in predicting organic food behavior and the major role of beliefs and attitude in forming buying intention. Given the relevant role in explaining consumer behavior, consumer beliefs need to be continually explored to take into account consumers' dynamic understanding of organic food (Massey et al., 2018). This study aims at exploring consumer beliefs regarding organic food. Following Ajzen (1991), we defined beliefs as all meanings associated with a product, its production process, and its consumption. Exploring these beliefs is complex due to the nature of organic food. Organic labeling is a process-related credence attribute (Lee & Hwang, 2016), meaning the regulation of the production process and its compliance cannot be easily verified by consumers. This leads to subjective and diverse beliefs about organic food regarding, for example, its safety, healthiness, and environmental friendliness (Fernqvist & Ekelund, 2014).

This study analyzed online comments gathered from news websites and forums because they particularly lend themselves to the exploration of consumer beliefs. The rise of different social media has enabled and sparked users' desires to share opinions publicly on online platforms (Ziegele, 2016), where consumers can write comments stating their beliefs about issues such as organic food. They do so of their own initiative (that is, without being asked by a researcher), revealing what matters to them. This shows which beliefs are on their mind and therefore salient and prevents the influence of the research process and physical/virtual presence of a researcher on stated beliefs (Branthwaite & Patterson, 2011). An additional plus to using online comments as data is that they are abundantly available at little cost.

Despite these numerous advantages, social media data, including online comments posted on news websites and forums, are rarely used to explore consumer beliefs about organic food. An exception is Olson (2017), who analyzed online comments for the persistence of certain beliefs following the release of the Stanford meta study on the health effects of organic food consumption. Unlike Olson (2017), our study aimed at uncovering the variety of beliefs rather than analyzing changes in selected beliefs. The goal of this study is to create a comprehensive inventory of organic food beliefs using online comments about organic food posted on the forums and discussion boards of news websites from German-speaking countries and the US. To this end, our chosen method was content analysis, an unobtrusive technique to extract meaning from textual data (Krippendorff, 2019). We constructed a category system of organic food beliefs by assigning textual units to different belief categories. Additionally, we counted the instances per category as an indicator of the salience of the beliefs. Obviously, our study reflects the beliefs expressed by people active online, which limits the scope of the study. However, a large and growing share of the population is represented on the Internet. In 2017, 87% of EU citizens (EUROSTAT, 2018) and 78% of US citizens (NTIA, 2018) were active Internet users. The proportion of social media users was 52% in Germany (Initiative D21 e.V., 2017) and 74% in the US (NTIA, 2018). Further, 23% of German online users write comments on news websites at least once a month (Ziegele et al., 2017), and 15% of Germans participate in online discussion forums (VuMA, 2018).

We chose to analyze online comments from people in German-speaking countries (Germany, Austria, and Switzerland, hereafter referred to as DACH) and the US because these countries are among the most important organic food markets globally. In terms of total retail sales, the US is the leading organic food market in the world, followed by Germany (2nd), Switzerland (8th), and Austria (11th) (FiBL, 2020). In 2016, organic food products' market share of total food sales was 5.1% in Germany, 7.9% in Austria, 8.4% in Switzerland (Organic Food Production Alliance [BÖLW], 2018), and 5.2% in the US (OTA, 2018).

In light of the constant development and growth of organic markets, our category system of organic food beliefs aimed to contribute to an up-to-date picture of consumers' understanding of organic food. The category system can be used in a variety of ways. First, the beliefs from the category system can readily be used by future consumer

research surveys eliciting organic food beliefs. The consideration of a comprehensive set of beliefs reduces the risk of neglecting salient consumer beliefs, which compromises the reliability of predictions about belief-driven market behavior (Hughner et al., 2007; Massey et al., 2018; Ajzen, 1991). Second, our category system can inform economic experimental research about relevant organic food beliefs. In economics, beliefs—largely neglected until recently—are recognized as an important component in the analysis of behavior (e.g., via choice experiments). Lusk et al. (2014) showed that a consumer might refrain from choosing organic alternatives not because s/he does not value an intact environment, but rather because s/he does not believe in the environmental benefit of organic food. Third, the category system of organic food beliefs can serve as input for further large-scale social media studies using automated text analysis methods. Finally, knowledge about consumers' understanding of organic food is relevant for marketing practitioners and policy-makers when designing and communicating marketing messages and policy measures. The remainder of this paper proceeds with a review of the literature on organic food beliefs. A description of our method is presented next, along with a reporting of the content analysis results. To conclude, we discuss the results and their implications for future research and practice.

3.1.2 Review of literature about organic food beliefs

Several literature reviews have summarized the vast body of empirical studies on the perception of organic food and motives for organic food consumption (Aertsens et al., 2009; Hemmerling et al., 2015; Hughner et al., 2007, Pearson et al., 2010; Rana & Paul, 2017; Schleenbecker & Hamm, 2013; Thøgersen et al., 2017). These reviews have identified many different organic food beliefs, which we have structured into four themes: A first theme comprises product-related beliefs regarding healthiness, food safety, price, taste, quality and appearance, nutritional value, naturalness, environmentally-friendly packaging, genetically modified organism (GMO) free, and availability. A second theme concerns beliefs related to the production of organic food, e.g., environmental consequences and animal welfare. A third theme concerns beliefs regarding the organic labeling and certification systems, as well as country-of-origin. A fourth theme concerns beliefs related to the role of organic agriculture in supporting local economies and food security, as well as beliefs related to fashionableness and nostalgia.

The extant empirical studies on organic food can be divided into exploratory and confirmatory studies. In the early 2000s, there was considerable exploratory research using focus groups and interviews to identify organic food beliefs (e.g., the UK-based studies of Harper and Makatouni (2002) and Padel and Foster (2005)). Seeking to mitigate the attitude-behavior gap regarding organic products, Aschemann-Witzel and Niebuhr Aagaard (2014) investigated the thoughts of consumers shopping in Danish retail outlets. But there has been little recent exploratory research that systematically describes current organic food beliefs and even less recent research focusing on the DACH countries and the US. Moreover, extant exploratory research mainly uses focus groups and interviews. The potential for analyzing textual data (e.g., content analysis of online comments) created at the consumer's own initiative remains largely untapped. The previously-mentioned study by Olson (2017) advocates for the use of online comments in research. He identified typical organic food beliefs and showed that online comments reveal beliefs that have been rarely considered in previous research, e.g., beliefs regarding the production scales of organic farming. Next to exploratory research, there have been abundant confirmatory studies on organic food, like surveys or experiments. These typically elicit consumers' agreement with a limited number of beliefs, usually no more than a dozen. For example, Moser (2016) used four beliefs in a study in Germany; Grankvist and Biel (2001) used six beliefs in a study in Sweden; and Nuttavuthisit and Thøgersen (2017) used eight beliefs in a study in Thailand. These confirmatory studies commonly elicited positively-phrased beliefs (e.g., 'organic food is healthier') and only seldom negatively-phrased beliefs (e.g., 'organic products offer no health effect'). Elicitation of positively- and negatively-phrased beliefs regarding the same theme was less common. An exception is the survey by Lee and Hwang (2016), in which they elicited 17 both positively- and negatively-phrased beliefs regarding food safety, eco-friendliness, price, quality, and value.

3.1.3 Material and methods

3.1.3.1 Data collection

We used news websites and forums as the sources of online comments for our study. Compared to social networks, we deemed news websites and forums to be more insightful sources for building a comprehensive inventory of organic food beliefs. First, as further

elaborated in Section 3.1.5.1, these sources use platform-specific and anonymous user profiles, meaning users may be less motivated by impression management, an issue especially present on social networks with personal followership, like Facebook or Twitter (Ziegele, 2016). Second, news websites and forums allow for long user comments with detailed belief statements. In contrast, Twitter tweets are limited to 280 (formerly 140) characters, and emotions are often expressed through emoticons rather than written language, as found in a previous study of food-related tweets (Vidal et al., 2015). Third, news websites and forums are freely accessible to the researcher, while Facebook profiles are mostly private (Steinert-Threkeld, 2018).

To identify the news websites and forums that would form the pool of relevant sources of this study, we followed the literature on text analysis of large electronic text databases and used relevance sampling. Relevance sampling is the strategy of choice when, as in the case of online discussion forums, an exhaustive register of all databases is lacking (Krippendorff, 2019). In essence, relevance sampling defines the criteria for sources to be relevant to the research goal of the given study. The goal of our study was to reflect the organic food beliefs of a broad user spectrum comprising both buyers and non-buyers of organic food in DACH⁴ and the US. We defined the following four relevance criteria, which had to be met by all sources: First, to cover a wide range of opinions, the sources had to be online sources not exclusively targeting sustainability or organic food issues. We used media statistics of the respective countries to locate such mainstream sources. For news websites, we located the 10 online outlets of the print press with the highest print coverage and the 10 online news services with highest online traffic. We excluded those news websites that did not offer discussion boards for their readers. For forums, we also used media statistics to select high-traffic online forums. However, there is no complete register of forums. Therefore, we conducted an extensive Web search and used cross-references to other forums. Second, from those pre-selected news websites and forums, we selected those that had publicly available discussion boards. Third, we selected those sources with boards that contained discussion on organic food. We used the search queries *organic AND food* and *organic AND agriculture* (US sources), *bio*

⁴ The data from the three German-speaking countries were sampled and analyzed jointly because of their similar organic food markets and integrated online media markets. Cross-border media usage was also visible in the comment data.

AND Lebensmittel and *bio AND Landwirtschaft* (DACH sources). We selected all threads containing these keywords. The threads comprised the series of comments posted under an initial forum question or a news article. Fourth, we selected all threads on organic food published between 2007 and 2017 and publicly available as of November 2017. Choosing such an extended timeframe prevented the discussions from being dominated by single topics, such as food safety scandals. The pool of sources meeting these relevance criteria are listed in Table 2. All relevant threads from this pool of sources constitute the population from which we sampled in the subsequent section.

Table 2 Pool of relevant sources.

	DACH		USA
News websites	aargauerzeitung.ch br.de focus.de derstandard.at diepresse.at freitag.de handelsblatt.com heise.de kurier.at merkur.de nachrichten.at	nzz.ch rp-online.de spiegel.de srf.ch tagesanzeiger.ch tagesschau.de tagesspiegel.de taz.de waz.de welt.de zeit.de	chicagotribune.com huffingtonpost.com latimes.com nypost.com nytimes.com usatoday.com washingtonpost.com
Forums	bfriends.brigitte.de chefkoch.de glamour.de gutefrage.net kleiderkreisel.de	paradisi.de reddit.com webkoch.de wer-weiss-was.de	reddit.com usmessageboard.com disqus.com quora.com

Source: Own table.

3.1.3.2 Data sampling

We randomly sampled threads from the population of threads described above. We sampled threads instead of single comments because threads retain the context in which comments are written. Retaining context is important for a thorough analysis of beliefs.

Textual sampling differs from statistical sampling as applied in survey research (Krippendorff, 2019). Unlike in survey research, the sampling units (threads) are not created for the purpose of being analyzed. Thus, threads are not “equally informative” in the sense that the amount of content related to the research question and the level of detail differ. Further, the sampling units constrain the analysis units (here, belief statements), unlike in survey research where sampling and analysis units are the same. Sampling units cannot be treated as single observations when determining textual sample sizes. Given these differences, rather than aiming at samples with the same distributional properties as

the population, textual sampling aims at yielding a sample that represents the phenomena (here, organic food beliefs) that lead to answering the research question (Krippendorff, 2019). Our sampling strategy aimed at a sample size that represented the variety of organic food beliefs.⁵ To this end, we took the following random samples. First, we randomly sampled threads for a DACH pilot sample (five threads totaling 100 comments), which only served training and preparatory purposes as described in step 2 of the analysis. Second, we randomly sampled threads for the DACH sample (60 threads totaling 1,094 comments). The random sampling resulted in approximately 70% of comments stemming from news websites and 30% from forums. This reflected the fact that in both the DACH and the US population of threads, there were more comments on news websites than on forums. Third, we randomly sampled threads for the US sample (47 threads totaling 1,069 comments), subject to the restriction of yielding a similar number of comments and maintaining a constant ratio of news website and forum comments. On average, one thread consisted of 18 comments in the DACH sample and 25 comments in the US sample. The mean length of the comments was 62 words for DACH and 99 words for the US.⁶

3.1.3.3 Data analysis

For analysis, the data were imported into the software MAXQDA 12. The content analysis was conducted following the requirements of objectivity, systematization, and quantification (Kassarjian, 1977; Kolbe & Burnett, 1991). Objectivity was pursued by conducting the analysis with two trained coders, ensuring intercoder reliability. Systematization was established by developing and applying a category system, which is a central analysis tool and the result of any content analysis. Quantification was established by counting the frequencies of units of analysis per category. A global chi-square test evaluated significant differences in the category frequencies between DACH and the US. The chi-square test per cell indicated the source of variation of the global chi-square (Symoneaux et al., 2012).

⁵ Previous content analyses have used similar sample sizes and time spans, for example, Camacho-Otero et al. (2019); Brún et al. (2014), Cho and Hong (2009).

⁶ In both the DACH and the US sample, very few of the sampled threads had more than 120 comments. We removed these threads as the discussion in these threads shifted away from the topic of organic food.

The analysis proceeded in the following steps.

- (1) In analysis step 1, we identified the units of analysis in our data, namely, belief statements about organic food. That is, we searched for any reference to organic food or organic farming. An example of a belief statement from our sample is “I truly believe that nature benefits from organic farming.” We excluded the mentions of organic products as recipe ingredients, mentions of the word ‘organic’ in a posted URL, and mentions of ‘organic’ in legal texts copied into comments, as these do not correspond to our definition of a belief statement as a written expression of a person’s own perception of and opinion about organic food. In the DACH sample, we identified 739 belief statements, and in the US sample, we identified 1,065 belief statements.
- (2) In analysis step 2, we developed a category system using literature and the pilot sample. A preliminary category system was drafted from all beliefs and superordinate themes identified in the literature review. Using this preliminary category system, the two coders simultaneously categorized the belief statements identified in the pilot sample (step 1). For example, the belief statement “I truly believe nature benefits from organic farming” was categorized into the belief category ‘organic protects the environment’, and this belief category was attributed to the theme category ‘environment’. In parallel to applying the preliminary category system, the coders extended the system by formulating novel belief and theme categories for belief statements that could not be described using the categories found in the literature (Krippendorff, 2019). The two coders compared their categorizations, and for each category they formulated a definition, an example from our data, and a coding rule.
- (3) In analysis step 3, the category system was applied to the DACH sample. The two coders simultaneously categorized the first 9% of the DACH sample to ensure the reliability of the categorization process. The coders again compared results and made minor adaptations to the category system and coding. The intercoder reliability was 87% using Cohen’s Kappa (Lombard et al., 2002). Afterwards, the coders proceeded with the categorization of the remainder of the DACH sample.

- (4) In analysis step 4, the category system was applied to the US sample. The two coders again simultaneously categorized the first 9% of the US sample to ensure the reliability of the categorization process. The coders compared results, and it proved necessary to add several belief categories specific to the US context. Cohen's Kappa for intercoder reliability was 85%, so the adapted category system of organic food beliefs was applied to the remainder of the US sample.

3.1.4 Results

The content analysis resulted in a comprehensive category system, displayed in Table 3. First, we took stock of organic food beliefs in the DACH and US samples. The category system revealed a variety of 65 organic food beliefs structured into 22 superordinate themes. Many of the 65 beliefs and 22 themes had been referenced by previous literature. In addition, the analysis revealed a considerable number of beliefs never or very rarely mentioned in previous literature (denoted with † in Table 3). The 65 beliefs and 22 themes had a wide thematic breadth, e.g., spanning from 'price' to 'food security'. We further structured the 22 themes into four main themes: 'product', 'food system', 'authenticity', and 'production'. Second, we counted the frequencies of belief statements per category to show how salient the different beliefs, themes, and main themes were in the online discussion about organic food. The frequencies of belief statements per category differed significantly between DACH and the US, as revealed by the global chi-square (main theme categories: $\chi^2 = 139.7$, $p < .0001$; theme categories: $\chi^2 = 236.7$, $p < .0001$; belief categories: $\chi^2 = 416.0$, $p < .0001$). Table 3 reports the relative frequencies of belief statements per category and how the frequencies differ between DACH and the US. Table 3 and the following sections of this paper present the main themes and themes in the order of descending salience in the US sample. Further, beliefs within a theme are clustered by positive, neutral, and negative valence (denoted with P, NT, and N in Table 3).

Table 3 The category system of organic food beliefs. The relative frequencies of belief statements (DACH: n=739; US: n=1,065) per category (main theme, theme, belief).

Main Themes		Themes		Beliefs				
DACH	US	DACH	US	DACH	US			
Product	32.3	44.1 (+)***	Food safety	8.7 (-)*	14.0 (+)***	OPs contain no or less chemicals compared to conventional products ^P	5.3	6.3
						OPs are safer and better regulated than conventional products ^P	0.5	1.0
						OPs also contain chemicals ^N	1.2 (-)**	3.2
						Conventional food safety is sufficient to ensure health ^N	0.8 (-)*	2.3
						OPs are not safe ^N	0.8	1.2
			Price	9.1	8.0	OPs are worth the price premium ^P	0.7 (-)*	1.8
						OPs are not (much) more expensive than conventional products ^{P+}	2.4 (+)***	0.3 (-)***
						OPs are a luxury ^{NT}	0.9	0.7
						OPs are (more) expensive than conventional products ^N	3.9	2.7 (-)***
						OPs are not worth the price premium ^N	1.1 (-)*	2.5
Healthiness	4.5 (-)*	7.9 (+)**	OPs are healthier than conventional products ^P	2.7 (-)**	5.4			
Taste	4.6	4.9	OPs offer no health benefit compared to conventional products ^N	1.8	2.4			
			OPs taste good or better than conventional products ^P	2.4	3.9			
			There is no taste difference between OPs and conventional products ^{NT}	0.8	0.8			
			OPs taste bad or worse than conventional products ^N	1.4 (+)***	0.1 (-)***			
Nutritional value	0.5 (-)***	4.3 (+)***	OPs are more nutritious than conventional products ^P	0.3 (-)***	2.3 (+)*			
			OPs are not more nutritious than conventional products ^N	0.3 (-)**	2.0 (+)*			
GMO	0.4 (-)*	1.8 (+)**	OPs are GMO-free ^P	0.4 (-)**	1.8			
Quality	1.9	1.3	OPs are fresher or of better quality or appearance than conventional products ^P	1.1	1.1			
			OPs are not fresher or of better quality or appearance than conventional products ^N	0.8 (+)*	0.2 (-)**			
Naturalness	1.4	1.2	OPs are (more) natural than conventional products ^P	1.4	1.2			
Availability	1.4	0.8	OPs are easily available ^P	0.3	0.5			
			OPs are not easily available ^N	1.1 (+)*	0.3 (-)**			
Food System	15.2 (-)***	26.8 (+)***	System integrity	8.5 (-)***	20.3 (+)***	OF is the sustainable option for the future ^P	0.4 (-)**	2.3 (+)*
						OPs are ethically produced ^P	0.4 (-)*	1.7
						The conventional agro-industrial lobby harms the organic food system ^{P+}	0.7	1.4
						My feeling says organic is better ^P	1.2	1.2
						OF internalizes the true costs of production ^{P+}	0.4	0.5
						Organic is not the priority in global health & sustainability challenges ^{NT+}	0.3 (-)**	2.1 (+)*
						Best practice approach of organic and conventional farming ^{NT+}	0.1 (-)**	1.3
						Buying OPs says a lot about the person ^{NT}	2.4 (+)***	0.4 (-)***
						The OFS cannot be trusted, it is just marketing ^N	2.4 (-)***	6.3
						The OFS is unsustainable ^{N+}	0.0 (-)***	1.7 (+)**

Four essays – *Essay I: Organic food beliefs in online comments*

Main Themes		Themes		Beliefs		DACH	US
DACH	US	DACH	US	DACH	US		
		Food security	3.5	3.5	I distrust the organic lobbyism/leadership ^{N†} Organic is a fad ^{N†} OF yields can feed the world ^P OF yields cannot feed the world ^N	0.0 (-)* 0.1 1.4 2.2	0.8 0.7 0.8 (-)* 2.7
		Production scale	1.6	1.5	OF is small scale ^{NT†} I disapprove the large-scale organic industry ^{N†}	1.6 (+)* 0.0 (-)**	0.6 (-)** 0.9 (+)*
		Farmer welfare	1.5	1.5	OF is profitable ^{P†} Organic certification is (too) expensive for farmers ^{N†} OF is not profitable ^{N†}	0.0 (-)* 0.9 0.5	0.7 0.6 0.3
Authenticity	39.4 (+)** 15.6 (-)**	Organic labels	21.1 (+)** 8.3 (-)**	8.3 (-)**	OLs can be trusted ^P	2.4	1.6 (-)**
					These OLs/OL criteria are trustworthy/reasonable ^{P†}	5.4 (+)**	0.3 (-)**
					OL criteria are not strict enough ^{NT†}	2.4 (+)*	1.1 (-)**
					OLs cannot be trusted ^N	4.9 (+)*	3.0 (-)**
					The term organic has no meaning ^{N†}	0.0 (-)**	1.0 (+)*
					These OLs/OL criteria are not trustworthy/reasonable ^{N†}	3.5 (+)**	0.8 (-)**
					OLs are confusing ^N	2.4 (+)**	0.4 (-)**
Origin	7.8 (+)** 4.2 (-)**	4.2 (-)**	OPs of local origin are more organic ^{P†}	0.5	0.2 (-)*		
			Local origin is more/as important than/as an OLs ^{NT†}	5.5 (+)**	2.5 (-)**		
			OPs from certain countries-of-origin are more organic than from others ^{NT}	1.8	1.5		
			OPs from certain retail outlets/brands are more organic ^P	4.7 (+)**	0.8 (-)**		
Retail outlet/brand	7.8 (+)** 2.2 (-)**	2.2 (-)**	OPs from supermarkets or discounters are less organic ^N	3.1 (+)**	1.3 (-)**		
			The importance of organic depends on the product category ^{NT}	1.9	0.9 (-)**		
Product category	1.9	0.9	Organic produce wrapped in plastic is less organic ^{N†}	0.7 (+)**	0.0 (-)**		
Packaging	0.7 (+)**	0.0 (-)**					
Production	13.1	13.5	5.5 (-)** 10.6 (+)**	10.6 (+)**	OF protects the environment ^P	2.0	3.8
					Conventional farming harms the environment ^P	0.7 (-)**	3.4 (+)*
					OF uses no or less chemicals compared to conventional farming ^P	1.8	2.1
					OF harms the environment ^N	0.9	1.2
					Conventional farming sufficiently protects the environment ^N	0.1	0.2
					OF protects animal welfare ^P	3.9 (+)**	0.7 (-)**
					OF does not guarantee animal welfare ^{N†}	1.2	0.7 (-)*
Biodiversity	1.9 (+)*	0.7 (-)*	OF protects biodiversity ^{P†}	1.9 (+)*	0.7 (-)**		
Working conditions	0.5	0.9	Conventional farming harms farmers' health ^{N†}	0.5	0.9		

Source: own table.

Notes: OP (Organic products), OF (organic farming), OL (organic label), OFS (organic food system). † Beliefs that have not or hardly been elicited in previous confirmatory studies. Chi-square per cell test: (+) or (-) indicate that the observed frequency is higher or lower than the expected frequency, *p < 0.05; **p < 0.01; ***p < 0.001.

3.1.4.1 Beliefs regarding the product

A salient main theme was ‘product’ and subsumed the following themes: ‘food safety’, ‘price’, ‘healthiness’, ‘taste’, ‘nutritional value’, ‘GMO’, ‘quality’, ‘naturalness’, and ‘availability’. It comprised beliefs directly associated with the attributes of organic food products and the personal costs and benefits arising from their consumption. In the US sample, 44.1% of all belief statements referred to the main theme ‘product’. ‘Product’ was the most salient theme in the US and significantly more salient than in DACH. In DACH, 32.3% of all belief statements were attributed to this main theme, making it the second most salient DACH.

‘Food safety’ and ‘healthiness’ were among the most important ‘product’-related themes in both regions. This mirrored the wide consensus in the literature that food safety and healthiness are the most salient attributes of and purchasing reasons for organic food (Hemmerling et al., 2015; Hughner et al., 2007). Both ‘food safety’ (DACH: 8.7%, US: 14.0%) and ‘healthiness’ (DACH: 4.5%, US: 7.9%) were significantly more important in the US than in DACH. The most prominent ‘food safety’ belief in both regions was that ‘organic products contain no or less chemicals compared to conventional products’ (e.g., pesticides, fertilizers, antibiotics, and hormones). However, some comments—especially those from the US—also contained the contrasting beliefs that ‘organic products also contain chemicals’ and ‘conventional food safety is sufficient to ensure health’. Industrial production leads to ‘mass production of high quality’, as one DACH commenter put it. In both regions, results showed that organic foods are more often believed to be healthy than unhealthy. In the US, organic food was often characterized as the opposite of unhealthy and industrially-produced foods.

Another salient ‘product’ theme was ‘price’, which was the subject of 9.1% of DACH and 8.0% of US belief statements. Surveys often only elicit the ‘negative’ belief that organic food is more expensive (e.g., Lee & Hwang, 2016; Lee & Yun, 2015; Thøgersen et al., 2015). Indeed, the most prominent price belief in both samples was that ‘organic products are (more) expensive’. However, our analysis revealed that price beliefs are far more differentiated and complex. Two additional negative beliefs were that ‘organic products are not worth the price premium’ and ‘organic products are a luxury’ only afforded by the wealthy. These beliefs mirrored previous findings that name price as the

most significant purchasing barrier (Hemmerling et al., 2015, Hughner et al., 2007). In spite of that, our content analysis revealed two positive price beliefs, namely, ‘organic products are worth the price premium’ and ‘organic products are not (much) more expensive than conventional products’. On the one hand, positive price beliefs could be partly grounded in an unwillingness to complain publicly about price. On the other hand, positive price beliefs—although less frequent than negative price beliefs—are relevant because of the possible behavioral consequences. A consumer might find organic products expensive but buy them regardless because s/he feels they are worth the price.

‘Taste’ beliefs accounted for approximately 5% of belief statements in both regions and were mostly positive. The most prominent taste belief was that ‘organic products taste good or better than conventional products’. Previous studies found superior taste to be the second most important driver of organic purchases (Hemmerling et al., 2015, Hughner et al., 2007). While the belief that ‘there is no taste difference’ was rare, some DACH comments referred to the negative belief that ‘organic products taste bad or worse’.

The ‘nutritional value’ of organic products was discussed significantly more in the US sample (DACH: 0.5%, US: 4.3%). In both regions, there were equal shares of references to organic food having higher nutritional value compared to not. The belief that ‘organic products are GMO-free’ was significantly more present in the US sample (DACH: 0.4%, US: 1.8%). With GMO products far more ubiquitous in the US market, US commenters stated they would buy organic products to avoid genetically-modified products.

Little reference was made to the ‘quality’ of organic food (DACH: 1.9%, US: 1.3%), although organic products were believed to be ‘fresher or of better quality or appearance’ more often than not. In both regions, there was some mention of ‘naturalness’. Commenters believed that ‘organic products are (more) natural’ (DACH: 1.4%, US: 1.2%). Organic products were generally perceived as less processed and believed to contain fewer artificial ingredients than conventional products (e.g., colorants, preservatives, flavor enhancers). The ‘availability’ of organic food in retail was discussed very little in DACH and the US (DACH: 1.4%, US: 0.8%).

3.1.4.2 Beliefs regarding the food system

The main theme ‘food system’ comprised the following four themes: ‘system integrity’, ‘food security’, ‘production scale’, and ‘farmer welfare’. ‘Food system’ beliefs concerned

the need for a sustainable food system and involved discussion on the strengths and weaknesses of the organic and conventional food systems. Commenters conveyed that purchase decisions not only generate personal benefits but also mean opting for either of the two food systems. ‘Food system’ beliefs were significantly more present in the US sample, accounting for 26.8% of the belief statements, compared to only 15.2% in the DACH sample.

The most salient ‘food system’ theme was ‘system integrity’, which we defined as confidence in the trustworthiness and viability of the organic food system compared to the conventional food system. ‘System integrity’ was significantly more salient in the US sample (DACH: 8.5%, US: 20.3%), where it also was the most salient theme out of all 22 themes. In relation to ‘system integrity’, the analysis revealed particularly many beliefs not (or very rarely) described in previous literature. These beliefs were both in favor of and critical of the organic food system. The belief that ‘organic farming is the sustainable option for the future’ was particularly salient in the US sample. Some commenters also believed ‘organic products are ethically produced’ and that ‘organic farming internalizes the true costs of production’. Other commenters were less specific in expressing their positive beliefs about organic food, stating that their ‘feeling says organic is better’. US commenters in particular pointed out that buying organic food is not only for their own benefit: “it is not always about me”, as one US commenter put it. As for negative beliefs, a belief very prominent in the US sample was that ‘the organic food system cannot be trusted, it is just marketing’. Many US commenters used the phrase “organic is a scam” to express their distrust in the organic food system. One person said that “just like the low-fat myth, the ‘organic’ label is used by the food industry to make people think it's great for them and then to make lots of money.” Previous literature has stressed the importance of trust in the organic label and certification system (Nuttavuthisit & Thøgersen, 2017; Thorsøe, 2015). Additional negative beliefs included ‘the organic food system is unsustainable’ and ‘organic is a fad’. In the realm of ‘system integrity’, we also identified beliefs concerning different agents within the organic and conventional food systems. Some commenters criticized lobbyism stating that ‘the conventional agro-industrial lobby harms the organic food system’, while others despised ‘the organic lobbyism and leadership’. Commenters with positive beliefs about organic food suspected conspiracies in the conventional food system. In contrast, commenters voicing negative beliefs about organic food suspected conspiracies in the organic food system. There was

also the belief that ‘buying organic says a lot about the person’ with organic food buyers pictured as romantic idealists. This mirrors previous findings from personal interviews with US consumers, through which Zepeda and Deal (2009) found that organic shoppers were described by non-organic food shoppers with the negative stereotype ‘hippies’ and by light- and frequent-purchasers of organic foods with the rather positive stereotype ‘yuppies’. We suggest further research on the extent to which avoiding such negatively connoted stereotypes presents a purchase barrier for potential buyers. Finally, there were two ‘system integrity’ beliefs that testified to the fact that US commenters in particular were open to agricultural approaches other than organic and conventional agriculture. One belief was that ‘organic is not the priority in global health and sustainability challenges’; rather, the priority is obesity or climate change, *inter alia*. In particular, these commenters proposed changes in consumption patterns, such as switching to plant-based, fat- and sugar-reduced diets, as well as buying fewer processed foods and more foods grown locally. Another belief was that agro-systems of the future should pursue a ‘best practice approach of organic and conventional farming’, e.g., combining soil fertility measures and GMOs. One US commenter said that “its [organic] methods need to be considered alongside conventional one [sic!] to create the best balance of productivity and sustainability.” Such pragmatic or holistic beliefs have been rarely considered in previous literature.

A second ‘food system’ theme was ‘food security’. ‘Food security’ accounted for 3.5% of both DACH and US belief statements. ‘Food security’ beliefs concerned whether organic farming uses land resources efficiently and has the production capacity to assure global food security. In both samples, organic farming yields were more often believed to be unable to feed the world. As one DACH commenter put it: “Only scientifically based industrial food production will be able to feed the world. Anything else is a green illusion long refuted by reality.” Some of the commenters who did believe in the capacity of organic agriculture to ensure food security added that the latter is conditional on a simultaneous reduction of meat consumption.

Closely related to ‘food security’ was the theme ‘production scales’, referenced in 1.6% of DACH belief statements and 1.5% of US belief statements. However, these beliefs were opposing, reflecting the different organic markets and agricultural landscapes. DACH commenters believed that ‘organic production is small scale’. Commenters

nostalgically described organic food production as taking place on traditional, family-owned farms, which are the opposite of what is considered ‘factory farming’ or ‘mass production’, especially in the context of animal husbandry. Moreover, some DACH commenters considered buying organic products to support the local economy. In sharp contrast, US commenters viewed organic production as ‘big organic’ and ‘disapprove of the large-scale organic industry’. One US commenter said “organic is a \$100B/yr global business now. It's no longer simply a matter of buying from the local hippie instead of Big Capitalism.” Such ‘production scale’ beliefs were also revealed in Olson’s (2017) comment analysis, but they have yet to be elicited in survey research.

Only a small share of 1.5% of belief statements in both samples referred to ‘farmers’ economic welfare’. Commenters discussed whether organic farming was profitable in the long term, and they attributed lacking profitability and increased financial risks to the costs of conversion, certification, and a larger required workforce. Some commenters believed that ‘organic certification is (too) expensive for farmers’. Further, DACH commenters criticized the dependency on subsidies. Such ‘farmer welfare’ beliefs have not been considered in the previous literature.

3.1.4.3 Beliefs regarding authenticity

The beliefs regarding the main theme ‘authenticity’ related to the following five themes: ‘organic label’, ‘origin’, ‘retail outlet or brand’, ‘packaging’, and ‘product category’. The ‘authenticity’ beliefs concerned ‘how organic’ certain organic products are and the meaning of ‘true organic’. According to the commenters, the authenticity of an organic product depends not only on the compliance with the standards of a trusted organic label but also other cues, such as the region or country-of-origin; retail outlet or brand; product category; and packaging. Commenters judged organic products on a subjective continuum of authenticity, ranging from authentic to inauthentic, depending on what cues were perceived as indicators of authenticity. Borrowing from the definition of brand authenticity by Napoli et al. (2014), we defined organic authenticity as the subjective evaluation of genuineness—which could be termed ‘organic-ness’—that consumers ascribe to an organic product. In the DACH sample, ‘authenticity’ was the most discussed main theme, accounting for 39.4% of belief statements. ‘Authenticity’ had significantly higher salience in DACH. This stood in contrast to the US, where the authenticity of

organic products received relatively little attention, accounting for only 15.6% of belief statements.

The most salient ‘authenticity’ theme in both regions was ‘organic labels’. However, in the US, ‘organic labels’ had significantly lower salience (8.3%) compared to the DACH sample, where ‘organic labels’ accounted for 21.1% of all belief statements and was the most salient theme of all 22 themes. This is in line with previous literature that determined trust in labels as a decisive purchasing driver (Hughner et al., 2007). Many commenters were generally convinced that ‘organic labels cannot be trusted’. One DACH commenter stated that “labels are not always reliable. Therefore, not everything that is certified to be organic always contains organic”. Few commenters explicitly stated that, in general, ‘organic labels can be trusted’. However, our results revealed several label-related beliefs that go beyond whether the label in general can be trusted expressing uncertainty and criticism about labels. DACH commenters explicitly differentiated which ‘organic labels or label criteria are trustworthy/reasonable’ and which are not. DACH commenters deemed private labels, such as Demeter or Bioland, more trustworthy than the EU or German organic labels. Some DACH commenters found that the different ‘organic labels are confusing’. In the US sample, some commenters thought ‘the term organic has no meaning’, since any plant or animal product is technically organic. Labels other than the public USDA label were barely mentioned in the US sample, reflecting the reality that, in the US organic food market, other labels have little presence. Some commenters from both regions believed ‘organic label criteria are not strict enough’. For instance, one DACH commenter found the criteria arbitrary and preferred “a measure of resource use”.

After ‘organic labels’, the second-most salient ‘authenticity’ theme in both regions was ‘origin’, although significantly more so in the DACH sample (DACH: 7.8%, US: 4.2%). Commenters viewed organic production as closely connected with local production. Some commenters explicitly stated that ‘organic products of local origin are more organic’. Commenters expressed a perceived increase in control over food safety and label compliance for food that is both local and organic. Some stated they trusted organic products more if they knew the producers personally. In fact, it was frequently expressed that ‘local origin is more/as important than/as the organic labels’, particularly in the DACH sample. Local food was valued for its smaller environmental impact in terms of transportation. Our results reflected the previously detected trade-off between local and

organic food (e.g., Hasselbach & Roosen, 2015; Bazzani et al., 2017). Denver & Jensen (2014) found in an experiment with Danish consumers that individuals who recognized the benefits of organic products had a preference for both organic and local foods. In contrast, individuals recognizing the benefits of local food had a high preference for local food over organic. In addition to local ‘origin’, beliefs also related to country-of-origin, affecting the perceived authenticity of organic products. The commenters voiced a preference for domestic over imported products, confirming previous findings for both the EU (Feldmann & Hamm, 2015) and US (Xie et al., 2016). However, some US commenters preferred European products to US products, assuming they were higher quality and ‘truly organic’. For imported products, certain countries-of-origin were perceived as more ‘organic’ than others. Commenters expressed feelings of disappointment and betrayal in cases where they perceived a misfit between the organic label and a product’s country-of-origin. “Not sure about China's organic standards...” is how one US commenter put it. This skepticism might be the result of presumably poor certification controls in certain countries-of-origin. Commenters also felt the environmental impact of long-distance transportation of imported organic products lessened the environmental benefits. One DACH commenter felt that “due to transportation, organic bananas are as harmful to the environment as conventional ones”, a statement that suggests the authenticity cues sent by both the product’s organic label and country-of-origin can sometimes offset one another.

The third ‘authenticity’ theme was ‘retail outlet or brand’. This theme was significantly more salient in the DACH sample (DACH: 7.8%, US: 2.2%). Commenters believed ‘organic products from certain retail outlets or brands are more organic’. Farm shops, delivery boxes, and smaller retail outlets tended to be perceived as more ‘organic’ than larger ones. In both regions, commenters stated explicitly that ‘organic products from supermarkets or discounters are less organic’. As one DACH commenter put it: “What is offered as EU-organic in supermarkets or discount markets has little to do with organic.” Our results clearly suggest that retail outlet and brand considerably influence the perceived authenticity of organic products. However, the literature is only starting to discuss purchasing venue as a cue for organic authenticity. Ellison et al. (2016) found that US consumers were more confident that organic tomatoes were truly organic (and therefore willing to pay more for them) if sold at farmers markets or fresh format stores compared to supermarkets or supercenters.

The fourth ‘authenticity’ theme concerned ‘packaging’ and was only present in the DACH sample (0.7%). Commenters believed that the plastic packaging of organic fruits and vegetables made the produce appear less ‘organic’ to them. Although there are several studies on the packaging of organic food (for a review, see Hemmerling et al., 2015), none have explicitly looked at the choice of conventional produce over organic due to plastic packaging (a common practice in general retail to facilitate pricing).

Finally, the fifth ‘authenticity’ theme was ‘product category’ (DACH: 1.9%, US: 0.9%). Some commenters in both regions indicated that ‘the importance of organic depends on the product category’. Without further specification, they judged the organic label more important for fresh products (e.g., fruit, vegetables, meat, and eggs) and baby food and less important for processed and storable goods. This prioritization reflects actual sales shares of organic product categories, for example in Germany, where fresh products like eggs, fruit and vegetables, dairy and meat are pre-eminent (Organic Food Production Alliance [BÖLW], 2018).

3.1.4.4 Beliefs regarding production

The main theme ‘production’ comprised beliefs related to the following themes: ‘environment’, ‘animal welfare’, ‘biodiversity’, and ‘working conditions’. The ‘production’ beliefs concerned characteristics of the production process and the impact of the process on the environment, animals, and society. ‘Production’ was the least important main theme in both DACH and the US, referenced in only 13.1% and 13.5% of belief statements, respectively. Our results on the relatively low salience of production characteristics are consistent with previous findings that consumers are more familiar with organic products than organic farming practices and standards. For example, a 2002 study in the UK showed that consumers equate free-range eggs with organic eggs (Harper & Makatouni, 2002).

Within ‘production’, ‘environment’ was the most prominent theme. Environmental beliefs were significantly more present in US comments (DACH: 5.5%, US: 10.6%). Positive beliefs about organic farming were more frequent: ‘organic farming protects the environment’; ‘organic farming uses no or less chemicals’ (e.g., pesticides, fertilizers, antibiotics, hormones); and ‘conventional farming harms the environment’. Nevertheless, a small share of belief statements said that ‘conventional farming sufficiently protects the

environment’ and even ‘organic farming harms the environment’, with the use of copper sulfate in organic viticulture receiving special criticism.

While the theme ‘environment’ was more salient in the US comments, ‘animal welfare’ was significantly more salient in DACH comments; in fact, US commenters hardly touched the theme of ‘animal welfare’ (DACH: 5.1%, US: 1.3%). The respective belief was that ‘organic production protects animal welfare’. Nevertheless, some remarked that ‘organic farming does not guarantee animal welfare’.

Previous literature makes no mention of beliefs regarding ‘biodiversity’. However, commenters in both regions recognized that ‘organic farming protects biodiversity’ due to the wider practice of crop rotation and a ban on feed imports grown in deforested areas overseas. ‘Biodiversity’ was discussed significantly more in the DACH sample (DACH: 1.9%, US: 0.7%).

Previous organic food literature has only considered beliefs about health effects on consumers, but not on producers. However, some commenters commented on farmers’ ‘working conditions’ (DACH: 0.5%, US: 0.9%) and criticized the detrimental health effects of production of conventional foods by stating that ‘conventional farming harms farmers’ health’ (e.g., due to the exposure to pesticides).

3.1.5 Discussion

Our content analysis revealed a variety of 65 organic food beliefs, subsumed into 22 themes and four main themes. In the comments, we found organic food beliefs similar to those identified by previous literature (see literature review). That previously identified beliefs were present in the online comments confirms the representativeness of our samples in terms of the phenomena being studied and validates our results. In addition, our analysis uncovered novel beliefs concerning themes of price, biodiversity, animal welfare, working conditions, farmer welfare, organic label, origin, packaging, production scale, and system integrity.

The large variety of beliefs illustrates that commenters in both regions have a subjective and heterogeneous understanding of organic food. Commenters not only held beliefs about commonly-elicited product and production characteristics, but they also stated beliefs about the authenticity of organic products and sustainability of the organic versus conventional food systems. The salience of beliefs differed significantly between DACH

and the US across the four main themes. ‘Product beliefs’ were discussed frequently in both countries, though significantly more in the US, where ‘product’ was also the most important main theme. In both regions, ‘product’ beliefs extended beyond defined organic standards, with commenters emphasizing inferred qualities such as healthiness and taste. Vega-Zamora et al. (2013) obtained similar results when exploring Spanish consumers’ organic food associations using focus groups. Next, commenters discussed sustainability and viability of organic versus conventional ‘food systems’. ‘Food system’ beliefs were significantly more salient in the US sample. The ‘food system’ beliefs showed that commenters not only pondered the advantages and disadvantages of the consumption of organic versus conventional products but also the respective food systems. Discussion about the ‘authenticity’ of organic products was significantly more frequent in the DACH sample, where it was also the most salient main theme. Our findings show that authenticity of organic food products depends not merely on the organic label; commenters sought further authenticity cues originating from origin, retail outlet or brand, packaging, and product category. Previous choice experiments have revealed that if little information is available, the organic label has a halo effect. This means that organic products are chosen because of the label, despite poor performance in terms of other attributes (Evangelidis & van Osselaer, 2018). However, when more information on other attributes is available, the importance of the organic label decreases (Costanigro et al., 2015). In line with these previous studies, our findings suggest that consumers perceive a “continuum of authenticity” of organic products. Several cues contribute to the authenticity assessment. Beyond organic labels, these cues include product origin, retail outlet, brand, product category, and packaging. ‘Production’ was the least-discussed main theme in both regions. While US commenters were significantly more concerned with environmental outcomes, animal welfare was discussed substantially more by DACH commenters. These differences in the priorities of European and US consumers reflect previous results on food values in different cultural contexts (Bazzani et al., 2018).

3.1.5.1 Reflections on the method

We make a case for using online comments to explore consumer thinking. We suggest this method as a promising alternative to conventional interviews or focus groups for several reasons. First, our results show face validity as they reflect, and even extend, the beliefs found in previous studies. Future research should examine whether the beliefs

expressed by online commenters correspond to those of the general DACH and US population. Second, with its non-interventionist way of inquiry, content analysis of online comments proves especially suitable for detecting which beliefs are salient and determining how salience differs between the DACH and US regions. Third, content analysis of online comments uses data generated in a non-research context and thus avoids typical common method biases (Podsakoff et al., 2003). For example, social desirability bias can be the result of the presence of a researcher (Golder & Macy, 2014), and such a bias is particularly relevant to behaviors with ethical dimensions, such as organic food consumption (Bray et al., 2011; Cerri et al., 2019; Randall & Fernandes, 1991). However, social desirability bias is by no means absent in social media data. Users present themselves favorably to other users as a means of impression management (Hogan, 2010). A survey by Jacobsen et al. (2017) found that self-presentation and learning are the main motivations for participating in virtual food communities. Against this background, we chose online comments on news websites and forums where commenters are potentially less prone to impression management when using anonymous profiles (as opposed to social networks like Facebook and Twitter, where users have personal profiles and followers) (Ziegele, 2016). In conclusion, the content analysis of online comments proved to be a valuable and cost-effective tool for gaining insights into consumer thinking and exploring regional differences.

3.1.5.2 Research implications

Insights from the content analysis of user-generated content can serve as valuable input for future studies. On the one hand, the category system of organic food beliefs could be employed for large-scale automated text analysis. For instance, the category system could train supervised machine-learning algorithms to perform aspect-based sentiment analysis, enabling the analysis of all comments published in selected sources and during a specific timeframe. Further, repeat measurements taken in the future could track changes in beliefs after food scandals or policy decisions. There is a growing body of consumer research using automated text analysis. For methodological overviews see Berger et al. (2020), Hartmann et al. (2019), Humphreys et al. (2018), and Vidal et al. (2018); for empirical studies see Christensen et al. (2017), Ordenes et al. (2017), and Peschel et al. (2019).

On the other hand, this research can inform experimental and survey research about beliefs currently held by consumers. Our findings suggest that future confirmatory studies eliciting beliefs should include a larger variety to improve measurement of beliefs and predictability of behavior. Most studies elicit beliefs regarding primarily product attributes. However, beliefs regarding authenticity and the global food system might superpose typical beliefs regarding product attributes when it comes to the formation of attitudes. For instance, a consumer might not believe in the health benefits of eating organic food but still maintain a positive attitude toward organic food because s/he strongly believes that the organic food system is the more sustainable option for the future. This means the latter belief is more important and has more weight when evaluating beliefs and forming attitudes. Thus, consumers can make trade-offs between normative and non-normative beliefs (Lusk & Norwood, 2009).

The variety of beliefs identified in this paper are the result of considering a wider range of themes and the subtle differences in the phrasings of beliefs regarding the same theme. The 65 beliefs in the category system reflect the natural phrasing of consumers. Eisner et al. (2019) have shown how using respondents' own terminology to construct survey items improves the reliability and validity of attitude measurement. Our content analysis provides the following insights: First, previous literature has mostly elicited positively-phrased beliefs (e.g., 'organic products are healthier'). The literature tends to consider some themes to be drivers of organic food consumption (e.g., healthiness) and others to be barriers (e.g., price) (Hughner et al., 2007). However, the online comments revealed positively-, neutrally- and negatively-phrased beliefs regarding most themes (e.g., 'organic products offer no health benefit'). In fact, out of the 65 beliefs, there was a similar number of positively- and negatively-phrased belief categories. Thus, our analysis suggests that a theme, e.g., healthiness, can be a purchase driver for some and a barrier for others who do not believe in the health benefits of organic products. Further, we cannot conclude from a disagreement with a positively-phrased belief (e.g., 'organic farming protects the environment') that a person believes the opposite (e.g., 'organic farming harms the environment'). S/he might simply not expect positive effects (e.g., environmental protection). Therefore, we suggest that future confirmatory research elicits positive, neutral, and negative phrasings of beliefs.

Second, the category system of organic food beliefs comprised several beliefs regarding non-organic products, production, and food systems. The comments suggested that individuals derive beliefs regarding organic food from their beliefs regarding conventional food and vice versa, depending on which food system they feel more informed about and/or prefer. For instance, positive beliefs regarding organic animal husbandry arose from negative beliefs about conventional animal husbandry ('mass production'). Consistent with findings of Vega-Zamora et al. (2013), individuals stating positive beliefs viewed the organic label as a cue for superiority over conventional products. In contrast, commenters stating negative beliefs did not state they were particularly convinced by the conventional food system, but rather that they did not believe in the benefits of the organic food system. Thus, if we want to understand consumer decision-making in relation to organic food, beliefs regarding organic products and beliefs regarding conventional alternatives should be considered.

Third, the consideration of simultaneously-held beliefs regarding a theme could be decisive when it comes to truthfully determining the influence of consumers' beliefs on behavior. For example, some commenters thought that organic products were expensive. Nevertheless, they believed they are worth the price. Likewise, a consumer could be confident that organic products are safer than conventional products but also believe that conventional food safety is sufficient. In both examples, eliciting only the first belief would result in biased conclusions about the individual's attitude. Uncovering simultaneously-held beliefs regarding both organic and non-organic options and their competing evaluation is crucial, since these jointly influence attitude and can either reinforce one another or cancel one another out. To conclude, considering all salient beliefs and their evaluation can contribute to the validity of predicted behavior in future confirmatory research.

Finally, the analysis of online comments can be used as a market-monitoring tool with practical implications for communication about organic food. In this paper, online comments revealed the themes consumers in DACH and the US are most concerned about. Organic food marketers should consider these different concerns in order to spread positive beliefs and disperse ambiguous and negative beliefs about organic food. Our study shows that confusion around the meaning of the term 'organic' persists, even though the issues surrounding this have long been known (Aarset et al., 2004). Several

beliefs are rather speculative, such as organic food being healthier or organic farming being small-scale. These kinds of subjective beliefs create a gap between expectations and reality; consumers come to expect something different from what organic products can actually deliver, creating the potential for consumers to feel deceived. Thorsøe (2015) describes unrealistic expectations as the primary challenge behind maintaining trust in the organic food system. Knowing consumers' actual expectations and beliefs regarding a product, i.e., the degree to which consumers believe in the benefits of a product, is crucial to the successful marketing of products and the development and introduction of new products to the market (Lusk et al., 2014).

In addition, knowing consumers' beliefs can help policy-makers and certification bodies improve communication about organic and other sustainability standards, thereby creating realistic consumer expectations and enhancing the comparability of standards across retail outlets and countries of origin. Moreover, introducing multiple levels of organic and sustainability standards could better meet consumers' heterogeneous expectations and create a market somewhere between organic and conventional (Janssen et al., 2009). Consumers could then make compromise decisions, e.g., a trade-off between animal welfare and price, as suggested by Jonge et al. (2015).

3.1.6 Conclusions

By means of a content analysis of online comments, this study deepened the understanding of consumers' beliefs regarding organic food in German-speaking countries and the US. The identified organic food beliefs mirrored and extended the beliefs described in previous literature. This proves that online comments can serve as an alternative data source for exploratory consumer research. The availability of online comments on other topics—potentially those more sensitive than organic food—depends on consumers' willingness to share their viewpoints online. Furthermore, our study showed that online comments are useful to determine significant differences in the salience of consumer beliefs in German-speaking countries and the US.

3.2 Combining content analysis and neural networks to analyze discussion topics in online comments about organic food (*Essay II*)⁷

Abstract

Consumers increasingly share their opinions about products in social media. However, the analysis of this user-generated content is limited either to small, in-depth qualitative analyses or to larger but often more superficial analyses based on word frequencies. Using the example of online comments about organic food, we investigate the relationship between qualitative analyses and latest deep neural networks in three steps. First, a qualitative content analysis defines a class system of opinions. Second, a pre-trained neural network, the Universal Sentence Encoder, analyzes semantic features for each class. Third, we show by manual inspection and descriptive statistics that these features match with the given class structure from our qualitative study. We conclude that semantic features from deep pre-trained neural networks have the potential to serve for the analysis of larger data sets, in our case on organic food. We exemplify a way to scale up sample size while maintaining the detail of class systems provided by qualitative content analyses. As the USE is pre-trained on many domains, it can be applied to different domains than organic food and support consumer and public opinion researchers as well as marketing practitioners in further uncovering the potential of insights from user-generated content.

⁷ *Essay II* has been published as Danner, H.; Hagerer, G.; Kasischke, F.; Groh, G. (2020): Combining content analysis and neural networks to analyze discussion topics in online comments about organic food. Paper presented and published in the Conference Proceedings of "3rd International Conference on Advanced Research Methods and Analytics", July 8-9, 2020, Valencia, Spain. DOI: 10.4995/CARMA2020.2020.11632. Adaptations to fit the format of this thesis have been made.

CRedit authorship contribution statement: **Hannah Danner:** Conceptualization, Methodology, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing, Visualization, Project administration, Supervision, Funding acquisition. **Gerhard Hagerer:** Conceptualization, Methodology, Software, Data Curation, Writing – original draft, Writing – review & editing, **Florian Kasischke:** Formal analysis, Investigation, Data Curation. **Georg Groh:** Conceptualization, Supervision.

Keywords: deep neural networks; natural language processing; consumer research; content analysis; social media; organic food.

3.2.1 Introduction

Novel communication technologies sparked the desire of users to publicly share opinions on online platforms (Ziegele et al., 2014). These developments provide an increasing amount of user-generated content, such as online user comments, which can be exploited by marketing and consumer research to gain insights into consumer thinking (Balducci & Marinova, 2018). Beginning with Kozinets' (2002) netnography of online communities, social scientists have increasingly analyzed textual user-generated content with established methods such as content analysis (Krippendorff, 2019). However, due to time and human resources required, such qualitative analyses are limited to small data samples. More recently, advances in automated text analysis and data collection enable consumer researchers to efficiently analyze larger data sets in a short amount of time and facilitate the detection of patterns, and compare measurements over time or between data sets. For an overview of methods see Berger et al., 2020. Frequently employed methods are dictionary-based approaches (e.g., LIWC, Linguistic Inquiry and Word Count, Tausczik & Pennebaker, 2010) relying on word frequencies. Researchers using automated text analysis have started to incorporate methods from the field of natural language processing (NLP, such as of data-mining, data-preprocessing, simple classifiers, and topic models (Latent Dirichlet Allocation, Blei, 2012) (for an overview see Vidal et al., 2018). However, to the best of our knowledge, there has been little research on how qualitative and NLP methods can be combined fruitfully. Latest advances in NLP are neural networks that account for the semantic context of words, i.e., word embeddings (Mikolov et al., 2013), or sentences, i.e., sentence embeddings (Cer et al., 2018). In this paper, we explore how such embeddings particularly lend themselves to be combined with qualitative text analysis by matching the analysis-depth of the latter with the scope of pre-trained sentence embeddings. In three steps, we present a novel approach for how a qualitative content analysis can be combined and enhanced with deep neural networks for semantic similarity.

We apply the approach to the case of organic food. Not only is a growing share of consumers aware of and buys organic food (Hemmerling et al., 2015)—making it an

increasingly important consumer research topic—, consumers also voice their opinions about organic food online (Danner & Menapace, 2020; Meza & Park, 2016; Olson, 2017). The analysis of online user-generated content can thus deliver valuable insights into which product attributes and related topics matter to consumers and what could be potential purchase drivers and barriers.

3.2.2 Methodology

In step 1 of our approach, a qualitative text analysis is conducted to develop a class system and manually classify a data set of interest. In Step 2, we use semantic features from pre-trained neural networks to investigate the semantic characteristics and the respective frequencies for each class. Step 3 presents criteria to combine results of both methods.

3.2.2.1 Step 1 – Qualitative analysis

To exemplify the approach, for step 1, we draw on a recent qualitative content analysis by Danner and Menapace (2020) of online comments about organic food. They manually extracted and classified consumer opinions (referred to as beliefs) about organic food to understand consumers' perception of organic. The authors collected 1069 online comments about organic food from high-coverage US news websites (e.g., nytimes.com, washingtonpost.com) and forums (e.g., reddit.com, quora.com). The 1069 comments consisted of 5510 sentences. Among these 5510 sentences, the two coders identified 1065 containing belief statements about organic food and subsequently classified those belief statements into 64 belief classes and 21 superordinate topics. For example, the sentence stated by a commenter *organic farming is better for nature* was attributed to the belief class *organic farming protects the environment*, which in turn was attributed to the topic class *environment*. By counting the frequencies of belief statements per category, the authors presented a detailed picture of topics salient to the online commenters in the data.

3.2.2.2 Step 2 – Universal Sentence Encoder

Using the same data and class system as in step 1, we find similar sentences for each class using the Universal Sentence Encoder (USE). USE is a recent advance in NLP and deep learning (Cer et al., 2018). Its architecture is based on the widely adopted Transformer architecture (Vaswani et al., 2017). USE is a deep neural network model pre-trained on

large scale text corpora from many domains. From there, the statistical knowledge in terms of generalizable, intermediate, semantic vector representations, which are also referred to as features or embeddings, can be used to quantify the semantics of specific domains, here organic food. USE works on sentence level providing sentence embeddings. The semantics of a given sentence are expressed by its vector representation. When compared to other sentences, the cosine similarity ranges between 1 (similar) to -1 (dissimilar).

We applied USE to automatically find semantically similar sentences for each of the 64 beliefs identified by Danner and Menapace (2020) (e.g., *organic farming protects the environment*) (Table 4). First, USE transformed each of the 64 beliefs and the 5510 sentences into an embedding. Second, USE measured the cosine similarity, i.e. the angular distance, between the embedding of each of the 64 beliefs (also referred to as seed sentences) and each of the 5510 sentences. When choosing a low threshold level for cosine similarity (i.e., the closer to -1), many sentences are considered as similar, whereas at high levels fewer sentences are considered as similar.

3.2.2.3 Step 3 – Evaluation

Eventually, we determine the appropriate level of semantic similarity, i.e., the respective cosine similarity threshold level which yields similar frequencies compared to the qualitative content analysis as reference. To this end, we inspect the thresholding results for cosine similarity levels from 0.7 to 0.84 based on the following criteria. (1) In the content analysis, 1065 sentences were relevant as in containing beliefs about organic food. A meaningful sentence filtering should yield a similar amount of relevant sentences. (2) The number of sentences assigned into the different classes should be similar for both methods. Therefore, we inspected the relative class frequencies and also calculated the Pearson correlation between the class frequencies for different cosine similarity levels. Figure 2 displays a trade-off between semantic similarity and class frequencies: the lower the cosine similarity (i.e., the less similar the sentences), the higher the correlation between the two methods. (3) Manual inspection should confirm the semantic cohesion between the manually and the automatically assigned sentences. Note that we performed the evaluation at topic level (21 topic classes) as the 64 belief classes are very detailed

and in part semantically too similar (e.g., *organic farming is better for the environment* and *conventional farming harms the environment*).

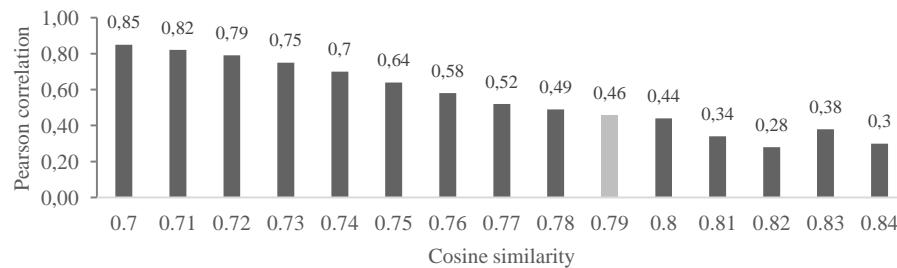


Figure 2 Pearson correlation of class frequencies (21 topic classes) between content analysis and USE. Source: Own illustration.

3.2.3 Results

Applying the aforementioned evaluation criteria, the thresholding performed best at a cosine similarity of 0.79. (1) At this level of similarity, USE found 1376 relevant sentences, which roughly corresponds to the 1065 relevant sentences identified in the manual analysis. (2) As highlighted in Figure 2, for cosine similarity of 0.79, both methods yielded similar class frequencies, indicated by a correlation of $r = 0.46$. However, class frequencies do not match perfectly. Looking at the relative class frequencies for each of the 21 topic classes in Figure 3, we find that the class frequencies for both methods are more similar for some topics than for others. For example, the topic *environment* accounts for 11% of sentences in the content analysis and 18% in the similarity thresholding. The most frequent topics in the content analysis were *system integrity*, *food safety*, *environment*; the most frequent topics in USE were *environment*, *system integrity*, *farmer welfare*. (3) For cosine similarity of 0.79, manual inspection showed very high semantic cohesion between the seed sentences per topic and the sentences identified as similar by USE. Table 4 displays the 11 sentences that USE found to be similar to the belief *organic farming protects the environment* at a cosine similarity of 0.79. All 11 are concerned with the effect of organic farming on the environment. However, sentences 3, 4, and 6 carry negative and thus the sentiment opposite to the seed sentence. Thus, while USE correctly identifies the topic, the sentiment is not always correctly classified, which is one reason why comparisons at topic level were chosen for this study. In addition, the manual inspection of the sentences classified by both methods proved that both methods classified largely the same sentences in the respective classes.

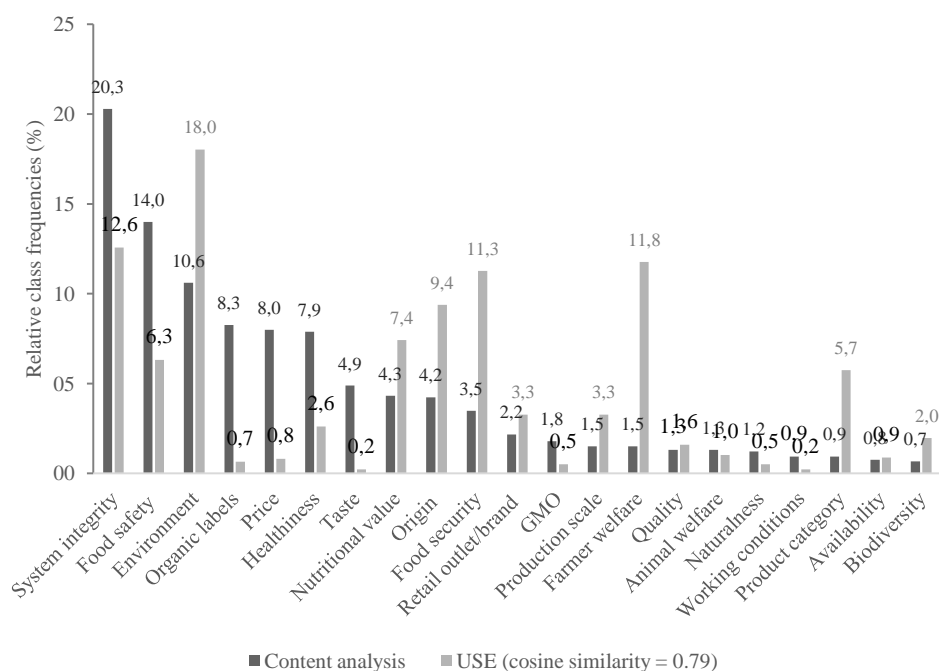


Figure 3 Relative class frequencies (21 topic classes) in content analysis and USE. Topics are ordered in descending frequency according to the content analysis. Source: Own illustration.

Table 4 A seed sentence from content analysis and the 11 sentences identified as similar by USE (cosine similarity = 0.79).

seed	Organic farming protects the environment.
1	Organic farming can help to preserve our environment for future generations.
2	The depletion of the soil and monoculture is what causes factory farming produce to be less nutritious than organic.
3	Mythbusting 101: Organic Farming > Conventional Agriculture
4	A lot of what I've read has said that organic farming is not better for the environment.
5	Organic is for the environment.
6	And from this we hear that organic farming is "devastating" to the environment.
7	Organic farming is much closer to the way Mother Nature farms.
8	GMOs can be super beneficial - to the consumer, the farmer, the environment.
9	Organic farming is greener
10	Besides delivering health benefits, organic farming is better for the environment.
11	Organic is for the environment.

Source: Own table.

3.2.4 Discussion

USE appears to be an effective and easy to use method to analyze large text corpora by searching for sentences that are semantically similar to seed sentences of interest. Seed sentences can originate, for instance, from a small-scale qualitative study—here the belief

classes identified by Danner & Menapace (2020). Provided a manually developed class system, it can analyze any unseen data set, —here 5510 sentences on organic food—, according to semantic similarity. In the present example, a human researcher selected the required level of similarity by evaluating the features generated by USE based on descriptive statistics and manual inspection. We suggested several criteria to select the appropriate similarity level as an alternative to training a classifier. Training a reliable classifier to classify fine-grained classes as complex as 64 different organic food beliefs requires large amounts of labeled data, which often exceed the resources of common research projects in the field of consumer and opinion research, and as it also applied to the presented example.

The selected similarity threshold was valid as the filtered sentences were widely coherent with the qualitative content analysis. In a subsequent step, USE could be applied to filter a larger unseen data set on organic food. Thus, the potential of the suggested approach lies in its scalability. We can extrapolate the detail of insight characteristic of qualitative research to analyze class frequencies in a larger data set of user-generated content.

Being still in an early phase, our approach bears potential for further refinements. We used a very large class system with 64 belief classes grouped into 21 topics, which also contained classes semantically very similar to each other. Using fewer and more distinct classes could thus improve the coherence between a manual classification and automatic classification based on USE. Furthermore, USE reliably finds the sentences containing similar topics, but does not always correctly distinguish positive and negative sentiment regarding the topic. Therefore, while suitable for topic classification, its use for sentiment analysis is bound to the manual control of a human researcher and domain expert. The imperfect match between manual classification and automatic filtering may also originate from the selection of the unit of analysis, a well-discussed issue in qualitative research (Campbell et al., 2013). The unit of analysis in USE are sentences, whereas in the content analysis, the unit of analysis could also stretch beyond a single sentence, and qualitative researchers can use domain knowledge for understanding and classifying text.

3.2.5 Conclusion

In a three-step approach, we suggested how a topic classification of a qualitative content analysis—here of online comments about organic food—can be combined with neural

networks like USE to find similar sentences. We proved that embedding techniques largely fit the results of qualitative analysis and point out their methodological potential. USE considers the semantic coherence between words and sentences and delivers in-depth insights by providing the original consumer phrasings (see Table 4) instead of abstract word lists and word frequencies as in more simple approaches of automated text analysis, such as dictionary-based approaches or LDA topic modeling.

Additional potential lies in cross-lingual applications using multilingual USE: Researchers can use the same seed sentences in one language and analyze data sets in different languages to make cross-country comparisons. Analyzing user-generated content, consumer researchers can learn about which product attributes and topics salient to consumers and potentially serve as purchase drivers or barriers. Based on this, consumer typologies and clusters can be derived. An improved understanding of consumers' opinions can support the design of organic products as well as labeling policies. Another application of USE lies in using items of established scales from survey research as seed sentences and analyze their similarity and prevalence in social media data. In addition, the suggested approach could be promising for market monitoring based on the targeted detection of social media content. For example, social media managers can observe the prevalence and development of certain opinions over time.

3.3 The news media and their audience: agenda setting on organic food in the United States and Germany (*Essay III*)⁸

Abstract

Agenda-setting research revealed that the news media are an important source of information and influence the topics salient in public opinion. This research investigates such agenda-setting effects between the news media and their audience in the context of organic food. A longitudinal text-mining study is conducted to analyze the salience of organic food topics in news articles and their reader comments published in major online news outlets in the United States and Germany from 2007 to 2020. Topics are modeled using a neural network approach based on clustered multilingual sentence embeddings. Findings show that the salience of topics in news articles significantly influences their salience in reader comments. However, reader comments have no effect on news articles. Metrics for agenda distance and agenda diversity confirm the media's agenda-setting role and additionally point out periods of time when events caused the media and public attention to diverge. The news media determine public opinion on organic food in the US and Germany and is thus an important player in the promotion of organic food consumption to be considered by marketers and policy makers.

Keywords: Agenda setting; media coverage; public opinion; organic food; text mining; topic modeling.

⁸ A similar version of *Essay III* is currently under review at the *Journal of Cleaner Production* as Danner, H.; Hagerer, G.; Pan, Y.; Groh, G.: The news media and their audience: agenda setting on organic food in the United States and Germany.

CRedit Authorship Contribution Statement: **Hannah Danner:** Conceptualization, Methodology, Validation, Formal analysis, Investigation, Writing - original draft, Writing – review & editing, Supervision, Visualization, Project administration, Funding acquisition. **Gerhard Hagerer:** Conceptualization, Methodology, Software, Formal analysis, Data curation, Investigation, Writing - original draft, Writing – review & editing, Visualization, Supervision. **Pan Yan:** Methodology, Software, Formal analysis, Data curation, Writing – original draft, Writing – review & editing, Visualization. **Georg Groh:** Conceptualization, Writing – review & editing, Supervision, Project administration, Funding acquisition.

3.3.1 Introduction

The news media are an important source of information. The Agenda-Setting Theory (McCombs & Shaw, 1972) postulates that the mass media influence the issues that become salient in public opinion. This theory laid the groundwork for a large body of research on how people's opinions are influenced by what they read in the media (Abdi-Herrle, 2018; Conway-Silva et al., 2018; Gerber et al., 2009; Hester & Gibson, 2003). Whereas the original agenda-setting hypothesis focused on the media determining *which issues* are salient (McCombs & Shaw, 1972), later research postulated that the media also successfully determine *how* an issue is framed and *which attributes* of an issue are salient (i.e., the thematic variety of an issue) and their relative salience. This has been conceptualized as attribute agenda-setting, second-level agenda-setting, or framing (Ghanem, 1997; McCombs, 2014; Wanta et al., 1995; Weaver, 2007).

Most agenda-setting research focuses on the political context (e.g., Conway-Silva et al., 2018; Gerber et al., 2009). However, there is yet little research investigating agenda-setting effects in a consumption context. If the media determine what is salient in consumers' minds, they can impact the formation of beliefs and attitudes toward consumption behaviors, and, thus, potentially play an important role in guiding desirable consumption patterns, such as buying environmentally friendly products. There is first evidence that the media influence consumers' environmental concern and attitudes, as antecedents of pro-environmental consumer behavior (Junsheng et al., 2019; Trivedi et al., 2018).

The environmental impact of households is particularly large with respect to food consumption (Tukker et al., 2010). In this regard, organic food is generally understood as a more sustainable alternative by both consumers and research (Siegrist et al., 2015; Thøgersen, 2010). Accordingly, organic food market shares have been rising in the past decades, reaching 5.7% of total food sales in the US and 5.3% in Germany in 2018 (Organic Food Production Alliance [BÖLW], 2020). The public perception of organic food has been elicited by an extensive body of literature via surveys, qualitative interviews, and focus groups (Hemmerling et al., 2015; Hughner et al., 2007). In parallel, organic food has made it on the agenda of the news media (Lockie, 2006; Meyers & Abrams, 2010) (Meyers and Abrams, 2010; Lockie, 2006).

However, thus far, news coverage on consumption issues and public opinion have been mostly analyzed separately. In particular, there is very little research on agenda setting in the context of organic food, that is, in how far news articles on organic food influence public opinion on organic food. An exception is Thøgersen (2006), who found first evidence for a connection between the media discourse and self-reported consumers' attitudes regarding organic food in Denmark. In addition, a new source of insight into public opinion on organic food has emerged in the meantime: user-generated content. Many news outlets offer comment sections granting their readers a platform to voice their opinions (Santana, 2011). Together with its increasing availability, such data is exploited by marketing and consumer research to gain insights into consumer thinking (Balducci & Marinova, 2018). In particular, the fact that users choose to comment online indicates their involvement with the issue and reveals what is naturally salient to consumers (Ksiazek et al., 2016). For instance, Danner and Menapace (2020) and Olson (2017) have detected organic food beliefs and topics and their salience in reader comments of US and German news media. However, such online user-generated content on organic food has not yet been exploited in the light of agenda setting.

In view of the outlined research gaps, the present research investigates the relationship between the salience of organic food topics in the news articles and comment sections of two major online US and German news outlets. Thus, this research investigates attribute agenda-setting, where the issue is *organic food* and the attributes are the topics associated with the issue of *organic food*. Using text mining, we analyze and compare the dynamics of the media and public agendas between January 2007 and February 2020, represented by the time-evolving relative distribution of topics in the news articles and comment sections.

The text data from news articles and reader comments are analyzed with a topic modeling approach based on clustered multilingual sentence embeddings. This technique from the field of natural language processing provides several novelties and advantages over classical approaches such as content analysis: Large amounts of data can be analyzed. Further, a topic modeling approach allows topics to emerge from the data without requiring prior knowledge. Moreover, the selected multilingual topic modeling approach enables the analysis and comparison of agendas in different countries and languages, here English and German. In consumer and communication research, there are as yet few

examples of studies using such an exploratory and data-driven topic modeling approach (Guo et al., 2016). An exception is the agenda-setting study by Pinto et al. (2019), who analyzed how topics in the Argentinian news media are reflected in Twitter activity and Google searches.

In the following, we outline the theoretical background of our study (section 3.3.2). Subsequently, we describe the study's methodology and introduce the employed agenda-setting metrics (section 3.3.3), before presenting the results on the different agenda-setting metrics (section 3.3.4). The paper concludes with a discussion of results (section 3.3.5) and conclusions (section 3.3.6).

3.3.2 Theoretical background

The vast body of agenda-setting research has been carried out in the context of political opinions. For instance, several authors have shown how the media influence public opinions on politics and voting behavior (Conway-Silva et al., 2018; Gerber et al., 2009; Groshek & Groshek, 2014). In contrast, research is still scarce and outdated with regard to agenda setting in the context of environmental behavior, (food) production and consumption, and, in particular, organic food. A causal relationship between media and public agenda was documented for the salience of societal issues such as environmental pollution (Ader, 1995). Mahlau (1999) analyzed the image of the agricultural sector in German print media from 1980 to 1994 and compared the facticity of the image in the media and in the population, observing that both media and public agenda have biased views on agriculture. An agenda-setting function of the media was not found as topics and biases differed in the media and public agendas. Evidence for framing was found in the context of organic food in Denmark in the late 1990s and early 2000s (Thøgersen, 2006). As negative articles had become more newsworthy, a slight decline in several psychological indicators concerning organic food was reported, indicating a potential connection between the media frame and consumer attitudes.

So while there is first evidence on agenda setting in terms of framing, attribute agenda setting—i.e., the relationship between the actual topics (attributes) discussed in the media and public agendas—has not yet been researched in the context of organic food. Insights on the media's impact on the public agenda would be insightful for organic food consumption research and the promotion of organic food sales in practice. In the attribute

agenda-setting process, salience (i.e., accessibility or top-of-mind awareness) is transferred from the media to the public. Additionally, the frequency with which the organic food topics are mentioned in the comment sections indicates the importance readers ascribe to these topics. In other words, by raising their voice on the issue of organic food and related topics, readers make a deliberate judgment about issue and topic (attribute) importance (Takeshita, 2006). Investigating the online salience is relevant as perceived issue importance or concern (e.g., concern about pesticide use in agriculture) can affect consumer behavior. For example, concern about social and environmental issues was shown to influence organic food purchases in Denmark (Thøgersen & Ölander, 2006). Bitsch et al. (2014) found that the media coverage of food safety incidents in the US and Germany affected consumer risk perception and impacted actual food purchase behavior. Thus, to shed light on the media's influence on consumer behavior, it is crucial to uncover how the media influence topic salience.

To investigate this relationship between the topic salience in the media and the public agendas, agenda-setting research has long matched content analyses of the media agenda with public opinion surveys. However, the digitalization of the media landscape has left its imprint. Together with media content and readership, agenda-setting research has taken the shift from print to online news media. Nowadays, online comment boards of news outlets and other social media platforms provide new opportunities for insights into the public agenda and quantitative analyses thereof (Neuman et al., 2014).

Furthermore, agenda-setting studies up until the early 2000s mostly documented a unidirectional causality of agenda setting. Subsequent research has shown that the influence can be reciprocal (Denham, 2010): the manifold possibilities to voice opinions online turned readers into senders (Santana, 2011). The media can be influenced by the public agenda, when journalists respond to public interests. This is referred to as agenda building or reverse agenda-setting (Denham, 2010; McCombs, 2014).

Against this background, this study aims at investigating the media and public agendas of organic food and their reciprocal relationship by analyzing articles and reader comments in two leading US and German online news media outlets. First, we hypothesize an attribute agenda-setting relationship between news articles and their reader comments. Second, we hypothesize that the media agenda predicts the topics discussed in reader comments of later articles. Third, we investigate whether the public

agenda also influences subsequent media agenda. Fourth, we expect to find differences with regard to agenda diversity and distance in the media and public agendas as well as between countries.

3.3.3 Methodology

3.3.3.1 Data

We analyzed the news articles and reader comments of two major news outlets representative of the German and the United States (US) context, respectively. We selected *spiegel.de* and *nytimes.com*, which are high-quality, general-purpose media outlets leading in terms of print coverage and online views in their countries (statista, 2019, 2020). They will hereafter be referred to as US and GER. We selected a broad time frame spanning January 2007 to February 2020 to thoroughly observe the dynamics of the media and public agendas on organic food. Articles and related comments on the issue of organic food were identified using the search terms *organic food* and *organic farming* and the German equivalents *Bio-Lebensmittel* and *Bio-Landwirtschaft*. In total, we considered 534 articles and 41,320 comments from the US for the years 2007 to 2020, and 568 articles and 63,379 comments from Germany for the years 2007 to 2017 and the year 2020. Due to website restructuring, no German comments were available for 2018 and 2019.

3.3.3.2 Topic modeling

Overview. Alongside with the growing availability of large amounts of text data from online media and social media platforms, big-data analysis techniques have provided an alternative or complement to traditional methods such as content analysis to efficiently analyze text data (Neuman et al., 2014). For an overview of text-mining methods in communication research, see Guo et al. (2016) and in consumer research, see (Berger et al., 2020). The most frequently employed automated text analysis approaches in the social sciences are dictionary-based techniques (e.g., LIWC, Tausczik & Pennebaker, 2010). Moreover, methods from computer science—more specifically natural language processing (NLP)—have been gradually introduced in the social sciences (Berger et al., 2020; Guo et al., 2016; Jacobi et al., 2015). Unsupervised topic modeling is frequently employed and allows for exploratory topic analysis without requiring prior knowledge.

The most prominent topic-modeling approach is the Latent Dirichlet Allocation (LDA) established by Blei et al. (2003), which generates a probabilistic distribution of words and topics in text documents. However, NLP also offers more advanced topic modeling approaches that are based on deep neural networks and account for the semantic context of words via word embeddings (Mikolov et al.), or of sentences via sentence embeddings (Cer et al., 2018). This text-mining study employed a topic modeling approach based on multilingual sentence embeddings clustered with k-means clustering. We explain its foundations and implementation in the subsequent paragraphs.

Sentence Vectors. Recent advances in NLP are subject to the evolutions of deep learning (Devlin et al., 2018). One of the past years' most influential developments has been the development of pre-trained word embeddings (Le & Mikolov; Mikolov et al.), where each word is represented as a vector, which captures semantic and syntactic features as well as the context of a word. Similar words have close and different words distant vector representations. Knowledge regarding semantic similarity is derived from large-scale text corpora, e.g., Wikipedia, using unsupervised learning. Pre-trained word embeddings improve the performance on many NLP tasks compared to traditional NLP methods (Hossain et al., 2019). Further improvements are achieved by contextualized word embeddings (Devlin et al., 2018) and transformer networks (Yang et al., 2019), which constitute the state-of-the-art in NLP. This also holds true for sentence embeddings employed in the present study, as these are derived from word embeddings (Devlin et al., 2018).

The present study targets a multilingual environment in English and German. We therefore drew on Google's XLING model, which allows coherently embed both English and German texts, as it is additionally trained on respective translations (Chidambaram et al., 2019). It is accessible through *Tensorflow Hub* (Yang et al., 2018). An input sentence (German or English) is transformed into a 512-dimensional vector as shown in Figure 4. For two sentences in German or English with a similar meaning, the scalar product of the two sentence vectors approaches 1, whereas it becomes -1 in case of semantic dissimilarity—see Figure 4 on the right.

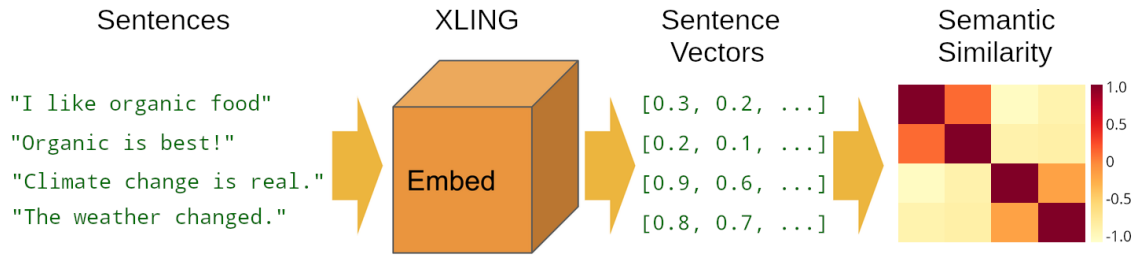


Figure 4 Sentence embeddings. Illustrative example of how sentence embeddings are generated using XLING. Inspired by Cer et al. (2018) and Shrimali (2018).

Topic Distributions. In NLP, the bag-of-words technique for document representation has a long tradition (Harris, 1954; Hossain et al., 2019). The occurrence of words is counted resulting in a histogram with the dimensionality of the whole vocabulary. Analogously, the bag-of-concepts technique (Kim et al., 2017) counts word or sentence vectors (Schmitt & Schuller, 2017), which can be clustered using different clustering techniques to model topics. The topic distribution of a document is thus a histogram representing the number of vectors in each cluster—see Figure 5. There are two approaches to bag-of-concepts in topic modeling. Sridhar and Kumar (2015) used *word2vec* word embeddings and Gaussian mixture models to derive the topic distributions based on soft quantization. He et al. (2017) proposed attention-based aspect extraction, i.e., an unsupervised algorithm using *word2vec* to derive meaningful semantic clusters from sentence vectors. In terms of topic coherence, both approaches have clearly outperformed traditional topic modeling such as Latent Dirichlet Allocation (Blei et al., 2003) not only for short texts such as Tweets or product reviews, but also for long texts such as scientific and news articles (Kim et al., 2017; Xing et al., 2014). For modeling the topics in the present study, we therefore opted for clustering pre-trained sentence embeddings instead of traditional methods such as LDA.

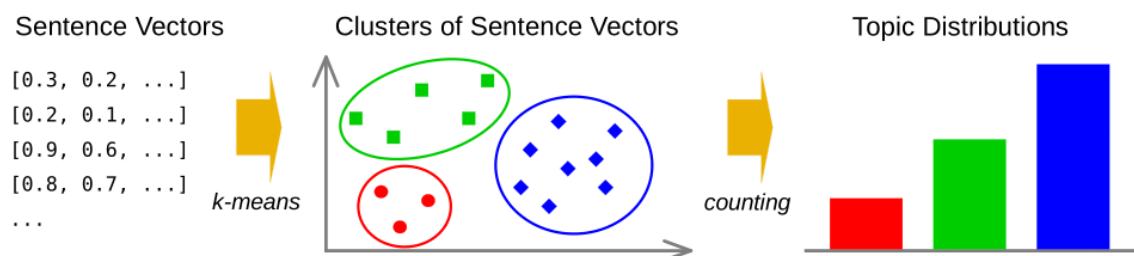


Figure 5 Topic modeling based on clustering of pre-trained sentence embeddings. Topic distributions are calculated for all sentences of a document. Here, a document is either a news article or the aggregated comments of a news article. Source: Own illustration.

Implementation. The topic modeling through the clustering of pre-trained sentence embeddings pursued the goal to develop a universal topic model for all data from US and GER articles and comments. First, we calculated a sentence vector for each available sentence in our data set using the pre-trained deep neural network XLING—Figure 4. Second, the resulting sentence vectors were clustered into topics using k-means clustering—see Figure 5. We derived the topic distribution per document (i.e., a news article or the aggregated comments of a news article) by counting the vectors in each cluster, yielding a histogram for each document as depicted in Figure 5. This means each document was not uniquely assigned to one topic but may realistically consist of different topics in different proportions. For the selection of the appropriate number of clusters k , i.e., topics, the Akaike Information Criterion indicated that a model with $k = 15$ was the optimal, parsimonious model. This was confirmed by a domain expert, who evaluated the semantic coherence of the topics based on the most important words (i.e., top words) representing each cluster. In the $k = 15$ model, five clusters were garbage clusters containing semantically incoherent top words. These clusters were excluded from further analysis, resulting in ten topics to be presented in the agenda-setting analyses. Third, the resulting topic clusters were labeled manually based on the top words (see Table 5 and Table 6). Two domain experts initially labeled the topics independently. There were only minor differences in the assigned topic labels, which were subsequently resolved by discussion.

Table 5 The most representative words (top words) for each topic in the US media and public agendas.

Topic Name	US Top Words
Food Safety & Chemicals & GMO	gmo, chemical, cancer, pesticide, organic, study, antibiotic, safe, fda, cause
Food Products & Quality	cheese, bread, cook, taste, tomato, recipe, bean, sauce, salad, fresh
Health & Nutrition	food, eat, diet, healthy, fat, sugar, nutrition, health, calorie, lunch
Environment & Climate Change & Energy	water, energy, carbon, climate, gas, heat, fuel, warming, air, emission
Farming	farmer, organic, crop, farming, agriculture, grow, soil, plant, pesticide, land
Animal Welfare & Meat Consumption	meat, animal, eat, vegetarian, vegan, beef, cow, chicken, feed, kill
Retailers & Prices	store, amazon, grocery, price, shop, company, product, market, customer, sell
Economy & Affordability	money, pay, tax, cost, profit, government, rich, school, care, income
Politics	trump, vote, political, government, republicans, president, party, conservative, democratic, obama
Evidence	science, article, study, read, fact, belief, cultural, religion, truth, evidence

Source: Own table.

Table 6 The most representative words (top words) for each topic in the German media and public agendas.

Topic Name	GER Top Words
Food Safety & Chemicals & GMO	dioxin, bio, bakterien, antibiotika, erreger, ehec, enthalten, gentechnik, gifte, krebs
Food Products & Quality	schmecken, käse, essen, kochen, brot, tomaten, milch, analogkäse, gurken, frisch
Health & Nutrition	lebensmittel, essen, ernährung, fett, gesund, nahrungsmittel, lebensmitteln, kalorien, zucker, ungesund
Environment & Climate Change & Energy	co2, erde, grad, energien, luft, wasser, erneuerbare, windkraft, atmosphäre, e10
Farming	landwirtschaft, bauern, bio, pflanzen, ökologisch, landwirte, anbau, konventionell, saatgut, dünger
Animal Welfare & Meat Consumption	fleisch, tiere, vegetarier, essen, hühner, massentierhaltung, kuh, fleischkonsum, futter, veganer
Retailers & Prices	verbraucher, kaufen, produkte, kunden, aldi, supermarkt, lebensmittel, produkt, qualität, ware
Economy & Affordability	geld, euro, zahlen, mittelschicht, steuern, bezahlen, einkommen, staat, kosten, mehr
Politics	spd, grün, politik, fdp, cdu, link, partei, merkel, deutsch, politiker
Evidence	verseuchung, menschen, schuld, aussagen, verstehen, lesen, diskussion, tragen, glauben, thema

Source: Own table.

3.3.3.3 Agenda-setting metrics

For the subsequent agenda-setting analyses, we evaluated and compared to which degree the ten identified topics were represented in the US and GER media and public agendas.

Hence, the unit of measurement applied in this research is the relative topic distribution. The proportion of a topic in the topic distribution of an article or its comments represents its salience, i.e., the degree of attention paid to the topic. To investigate agenda setting, we compared the relative topic distributions in the news articles and the comments using different metrics, which are outlined in Table 7. In addition to correlational analyses between single topic proportions—as they are typical of agenda-setting research—we applied two metrics from information and probability theory, which consider the distribution of all topics simultaneously: Normalized Shannon entropy for agenda diversity, and Jensen-Shannon distance for agenda distance. Further details can be found in the appendices. In spite of the lack of German comment data for the years 2018 and 2019, we opted to consider the time span up until 2020 to yield an up-to-date picture for the US articles and comments. For Germany, we still provided insights into the agenda diversity in articles up to 2020. However, agenda distance could consequently not be measured for the two missing years, and there were fewer correlational data pairs for the two correlation-based measures.

Table 7 The four agenda-setting metrics and how they measure the relationship between the media and public agendas.

Metric	Relationship between Media and Public Agendas
Pearson Correlation ρ_t	Linear relationship between the relative proportions of a topic in an article and its comments
Cross-Lagged Correlations $P_{X_1Y_2}, P_{Y_1X_2}$	Linear relationships between the relative proportions of a topic in an article and the comments of future articles, as well as the comments of an article and future articles
Normalized Shannon Entropy H	Agenda diversity, i.e., the diffusion of the relative topic distribution in an agenda over time. Local minima indicate time periods of low agenda diversity, i.e., certain topics dominate due to specific events.
Jensen-Shannon Distance (JSD)	Agenda distance, i.e., the similarity between the relative distribution of topics in the media and public agenda. Local maxima indicate time periods where the distance between media and public agendas is high

Source: Own table.

3.3.4 Results

3.3.4.1 The topics

The multilingual topic clustering revealed that the US and German media and public agendas were represented by the following ten topics: *Food Safety & Chemicals & GMO*, *Food Products & Quality*, *Health & Nutrition*, *Environment & Climate Change & Energy*, *Farming*, *Animal Welfare & Meat Consumption*, *Retailers & Prices*, *Economy*

& *Affordability*, *Politics*, and *Evidence*. The multilingual topic clustering allowed to identify the ten overarching organic food discussion topics in both countries and their comparison in the subsequent analyses. At the same time, country-specific characteristics were retained. This can be seen from differing top words (Table 5 and Table 6). In particular, the topic *Politics* revealed the political parties and politicians of the respective countries. Further, the topic *Retailers & Prices* indicated which retailers are present in the US and German food markets.

Figure 6 gives an overview of the salience of the ten topics in US and German articles and comments. For example, the topic *Retailers & Prices* was represented on average with more than 20% in US articles, and with approximately 8% in US comments. We detected clear differences between the US and Germany and slight differences between the media and the public agendas within in each country. The US media covered in particular the topics of *Farming*, *Health & Nutrition*, *Food Products & Quality*, *Food Safety & Chemicals & GMO*, and *Retailers & Prices*. The public agenda in the US also prioritized these topics; additionally, the topic *Evidence* was salient, indicating that consumers discuss the trustworthiness of organic food and the origin and reliability of information.

The most salient topics in the German media were *Farming*, *Retailers & Prices*, *Economy & Affordability*, *Politics*, and *Food Products & Quality*. The topic overlap between media and public agenda was smaller for Germany indicating slightly different agenda priorities. Commenters were mostly concerned with *Economy & Affordability*, *Politics*, *Evidence*, *Animal Welfare & Meat Consumption*, and *Environment & Climate Change & Energy*.

Comparing the salient topics between the two countries, we found that US media and public agenda focused on safety-, health-, quality- and price-related aspects of organic food, i.e., topics concerning the personal benefits of buying organic food. In contrast, the German media and public agenda were concerned with external impact of organic food production, i.e., the consequences for the environment and animal welfare. Further, they discussed political and economic aspects of organic food as well as different retailers. Overall, in Germany, attention seems to be more evenly distributed between the ten topics.

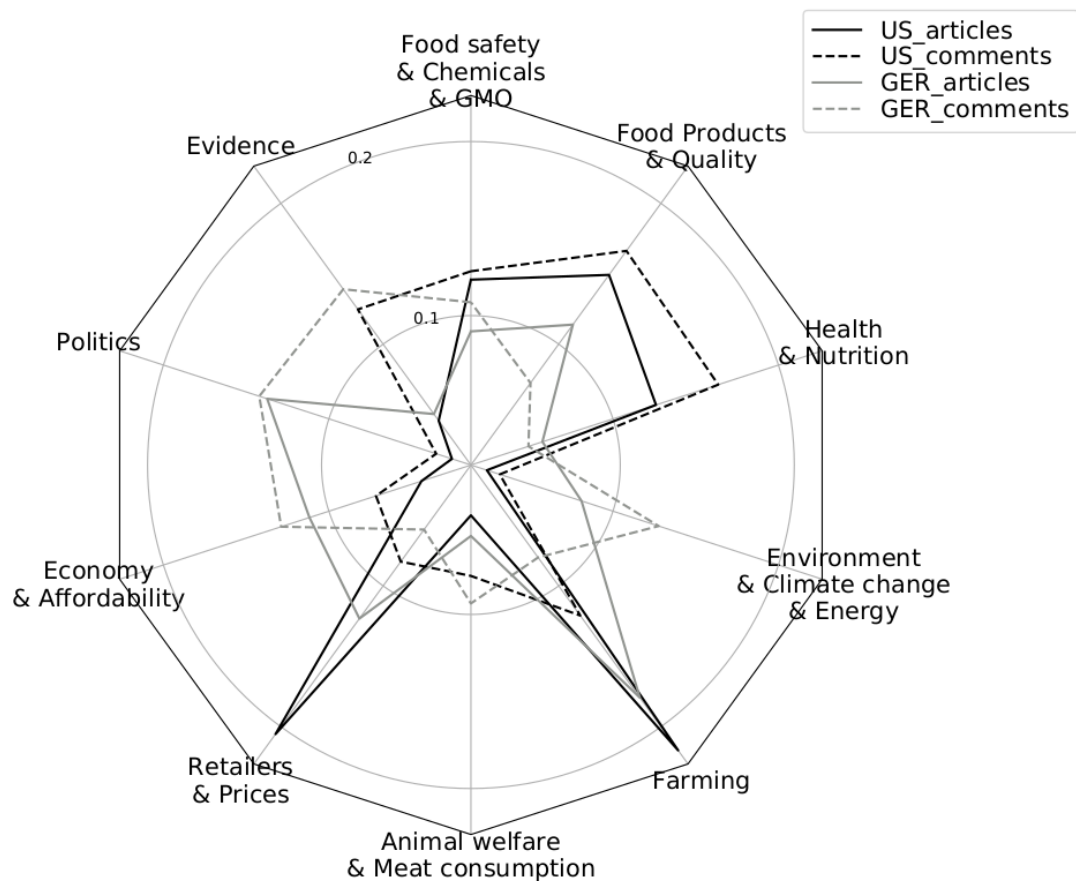


Figure 6 Average topic distributions (%) in US and GER media and public agenda. Source: Own illustration.

3.3.4.2 Synchronous topic correlations

In the following, we investigate the general agenda-setting function of the media. To that end, we determined the linear relationship between the relative proportion of a given topic in the articles and in the comments corresponding to each article by calculating the correlation between the two relative proportions. The coefficients ρ_i for the ten topics are displayed in Table 8.

We found a positive and significant linear correlation between each topic in an article and the same topic in the comments to that article for both US and Germany—see Table 8. In the US, all correlations were strong with coefficients ranging between $r = 0.540$ for *Evidence* and $r = 0.808$ for *Health & Nutrition*. For Germany, correlation coefficients range from a medium-sized correlation for evidence at $r = 0.423$ to strong correlations for all other topics, the strongest correlation being $r = 0.913$ for *Food Products & Quality*. All correlations were significant at $p < 0.001$. In conclusion, the strong associations found indicated that topic proportions in comments and articles were largely similar. This means

that the salience of any of the 10 topics in the article largely corresponded to the salience in the respective comments.

Table 8 Topic correlations ρ_t in media and public agenda for the US and Germany. All correlations are statistically significant ($p < 0.001$).

Topic	US	GER
Food Safety & Chemicals & GMO	0.712	0.773
Food Products & Quality	0.785	0.913
Health & Nutrition	0.808	0.690
Environment & Climate Change & Energy	0.626	0.764
Farming	0.759	0.754
Animal Welfare & Meat Consumption	0.781	0.757
Retailers & Prices	0.783	0.668
Economy & Affordability	0.679	0.765
Politics	0.711	0.817
Evidence	0.540	0.423
Total correlation	0.751	0.761

Source: Own table.

3.3.4.3 Cross-lagged topic correlations

In the previous section, we showed that the topics in the news articles strongly influence the topics of the respective comments. In this section, we investigate whether topics discussed in articles influenced not only their own comments but also comments of future articles P_{X1Y2} . In addition to agenda setting, we tested whether the topics discussed in comments at one point influenced the topics covered by the media in future articles P_{Y1X2} (reverse agenda-setting). To that end, we analyzed the cross-lagged correlations between a given topic in articles and comments (Kenny, 1975; Rogosa, 1980). We used a bi-weekly time lag, as both news outlets published an article on organic food on average every two weeks. To test the validity and direction of the cross-lagged correlation, we applied the Rozelle-Campbell Baseline (RCB) (Rozelle & Campbell, 1969) following the example of previous cross-lagged correlational analyses in the context of agenda setting (Abdi-Herrle, 2018; Lim, 2006). RCB designates the level of correlation to be expected on the basis of the autocorrelations and synchronous correlations. Valid cross-lagged correlations are present if they exceed the RCB.

First, total cross-lagged correlations indicated that the media agenda influenced the future public agenda, given $P_{X1Y2} > \text{RCB}$, in the US and Germany (see Figure 7). We found cross-lagged correlations (P_{X1Y2}) of $r = 0.375$ in the US and $r = 0.211$ in Germany. Second, the sizes of cross-lagged correlations differed between topics and countries when looking

at the topic-level correlations (see Table 9). In Germany, the topics *Environment & Climate Change & Energy* and *Animal Welfare & Meat Consumption* were correlated more strongly than other topics, indicating that the media had made these topics very salient, to the point that their salience in public opinion persisted in the subsequent time window. In the US, correlation sizes were more homogeneous across topics, with *Food Safety & Chemicals & GMO*, *Health & Nutrition*, and *Retailers & Prices* being the topics most correlated over time. Third, total cross-lagged correlations indicated that comments did not influence later media articles, given $P_{YIX2} < \text{RCB}$ (see Figure 7). For Germany, we hardly found any correlation for P_{YIX2} ($r = 0.008$). For the US, there was a weak correlation ($r = 0.158$), which, however, was not valid as it did not surpass the RCB. At topic-level none of the correlations between comments and articles was significant. For cross-lagged correlations, we conclude that the influence between articles and comments was unidirectional—articles influenced future comments, but comments did not influence future articles. Thus, there is no evidence for reverse agenda-setting as article authors did not seem to consider topics discussed in the comments. In addition, we found that topics salient to commenters in one time period maintained their salience in the subsequent time period. There were significant autocorrelations between the comments in consecutive time windows. For the articles, we found only a very weak indication for such agenda persistence in Germany ($r = 0.093$) and weak autocorrelations for the US ($r = 0.158$) (see Figure 7).

In conclusion, articles as well as comments from one time period influenced the topics discussed in later comments. In contrast, the media agenda was not susceptible to influence from previous articles or comments.

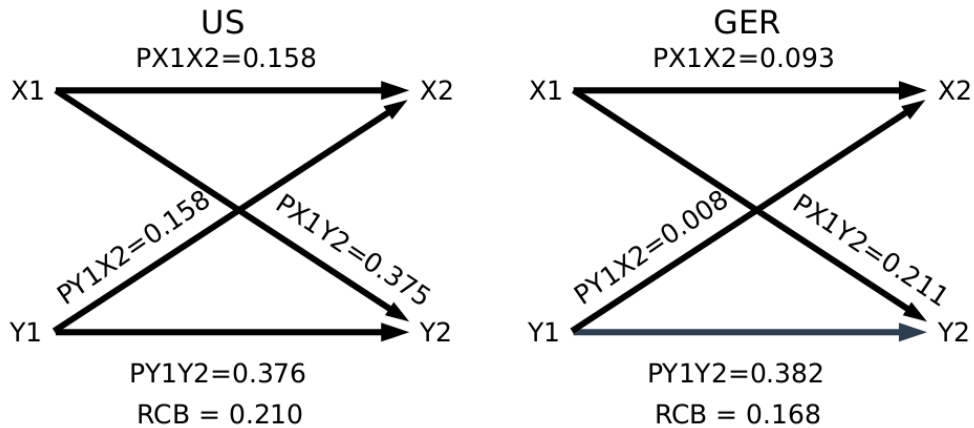


Figure 7 Total cross-lagged correlations. Cross-lagged correlations between articles and comments ($P_{X_1Y_2}$), and between comments and articles ($P_{Y_1X_2}$) of consecutive weeks for the US and Germany. Autocorrelations are given among articles ($P_{X_1X_2}$) and among comments ($P_{Y_1Y_2}$). All correlations are statistically significant at $p < 0.001$, except $P_{X_1Y_1}$ for Germany, which is significant at $p < 0.05$. Rozelle-Campbell Baseline (RCB). Own illustration based on Kenny (1975).

Table 9 Cross-lagged correlations between articles in t_1 and comments in t_2 ($P_{X_1Y_2}$), and between comments in t_1 and articles in t_2 ($P_{Y_1X_2}$). Rozelle-Campbell Baseline (RCB). Statistical significance is denoted with * ($p < 0.05$), ** ($p < 0.01$), and *** ($p < 0.001$).

Topic	US			GER		
	$P_{X_1Y_2}$	$P_{Y_1X_2}$	RCB	$P_{X_1Y_2}$	$P_{Y_1X_2}$	RCB
Food Safety & Chemicals & GMO	0.372 ***	0.105 ns	0.209	0.187 ***	0.012 ns	0.181
Food Products & Quality	0.206 *	0.023 ns	0.153	0.152 ***	-0.018 ns	0.230
Health & Nutrition	0.359 ***	0.119 ns	0.106	0.204 ***	-0.010 ns	0.137
Environment & Climate Change & Energy	0.180 ns	0.002 ns	0.101	0.566 ***	-0.013 ns	0.217
Farming	0.189 *	-0.157 ns	0.148	0.249 ***	0.040 ns	0.138
Animal Welfare & Meat Consumption	0.314 ***	-0.021 ns	0.124	0.402 ***	-0.022 ns	0.182
Retailers & Prices	0.340 ***	0.062 ns	0.216	0.220 ***	0.132 ns	0.097
Economy & Affordability	0.035 ns	-0.001 ns	0.026	0.055 ***	-0.154 ns	0.042
Politics	0.284 **	0.006 ns	0.162	0.267 ***	0.072 ns	0.190
Evidence	0.256 **	0.175 ns	0.033	0.003 ***	-0.016 ns	0.071
Total correlation	0.375 ***	0.158 ***	0.210	0.211 ***	0.008 ns	0.168

Source: Own table.

3.3.4.4 Agenda diversity

The previous sections provided evidence for agenda setting in the context of organic food. The salience of a topic in the media agenda predicted the salience of the topic in the public agenda. Subsequently, we used normalized Shannon entropy H to investigate how the diversity of topics in the media and public agenda evolved over time. H measures the diffusion of distributions. H increases when the topics are equally distributed, and H decreases when certain topics dominate. Thus, we can interpret time periods of low agenda diversity by identifying local minima of H . This answers the questions of whether topics are equally salient from 2007 to 2020, and whether certain topics dominate the

agenda in certain time periods. Figure 8 shows the agenda diversity H over time for the US and German media and public agendas. The insights from agenda diversity were two-fold.

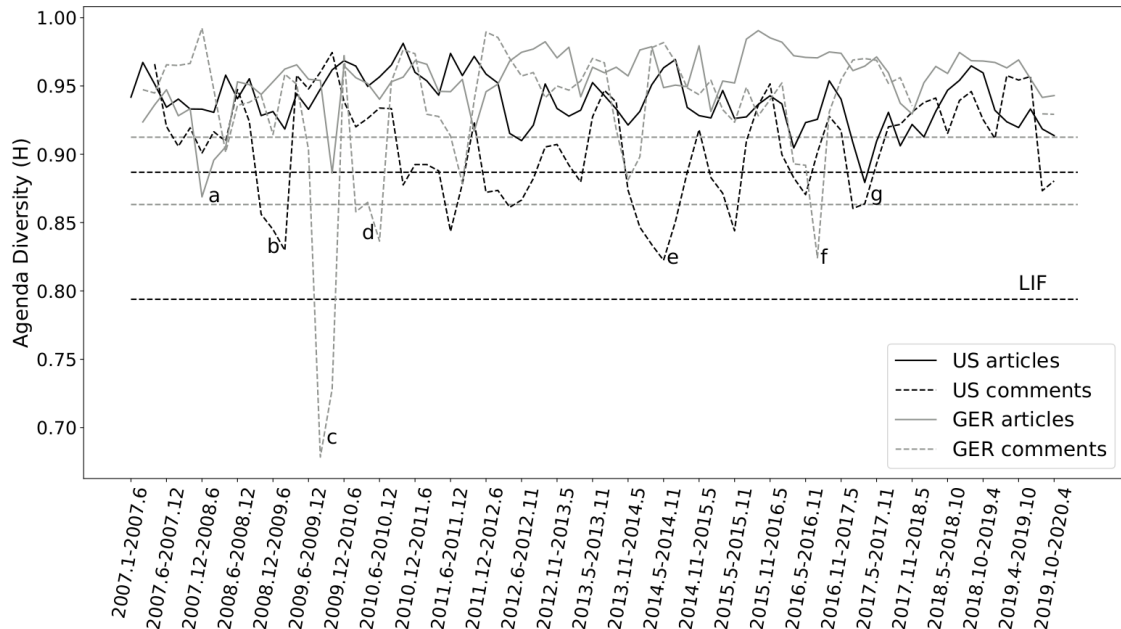


Figure 8 Agenda diversity over time. Measured by normalized Shannon entropy (H) for the US and German media and public agendas. The horizontal lines point out the lower inner fences (LIF) to identify local minima, i.e., time periods with low agenda diversity. These time periods are denoted with a , b , c , d , e , f , and g .

First, there were statistically significant differences in the levels of agenda diversity. In the US, we found that the diversity of articles was significantly higher than the diversity of comments ($F(1,154)=55.4$, $p < 0.001$). Likewise, in Germany, the articles were more diverse than the comments ($F(1,145) = 11.8$, $p < 0.01$). The observation that the media agenda was more diverse than the public agenda is consistent with our previous findings from topic correlations, which indicated that commenters stuck closely to the topics discussed in the articles. Across countries, the agenda of the German articles was significantly more diverse compared to those in the US ($F(1,156)=23.7$, $p < 0.001$). Also, the agenda of the German comments was more diverse than in the US ($F(1,143)=13.8$, $p < 0.001$). German commenters not only cover a wider range of topics, but also discussed more extensively: On average, German comments are more than twice as long than US comments in terms of sentences.

Second, the entropy metric let us identify time periods in which the agenda diversity is particularly low, i.e., where the agenda was dominated by specific topics. Between 2007

and 2020, we detected seven time periods with this type of agenda diversity minima, denoted with *a*, *b*, *c*, *d*, *e*, *f*, and *g* in Figure 8. For these time periods, the radar plots in Figure 9 show the topic distribution and thus which topic dominated in which agenda. By additionally inspecting the articles and comments from those time periods, the minima in agenda diversity could be linked to real events, leading to the following interpretations: Time period *a* coincides with a peak in global food prices, which were also connected to food riots in the advent of the Arab Spring. The German media agenda was dominated by the topics *Economy & Affordability* and *Farming* and discussed financial speculations with agricultural commodities, the strongly subsidized EU farming sector, and rising food prices against the background of global food security. Time period *b* covers the start of Barack Obama's first mandate as US president. The US public agenda was dominated by the topic *Health & Nutrition*, as commenters discuss the president's food policy as well as Michelle Obama promoting healthy nutrition and installing an organic vegetable garden on the White House premises. In time period *c*, the German public agenda showed a strong minimum. The public discussion was dominated by the topic *Environment & Climate Change & Energy* in light of the 2009 United Nations Climate Change Conference in Copenhagen. Additionally, commenters debated the environmental policy of the newly elected German conservative-liberal government. In time period *d*, the topics *Animal Welfare & Meat Consumption* and *Economy & Affordability* prevailed the German comments. Users discussed the introduction of the EU organic label in 2010 as well as animal welfare benefits of organic animal husbandry. The next minimum *e* was in 2014 with *Food Safety & Chemicals & GMO* as the predominant topic among US commenters. At the time, new scientific evidence on the food safety of GMO and organic food had been released and certain US states had introduced mandatory labeling of GMO-ingredients. In this context, users pondered the meaning of *naturalness*. In *f*, the topic *Politics* prevailed in the German public agenda given the federal elections of 2016. Finally, in time period *g*, we found a minimum in the US media agenda with *Retailers & Prices* being the dominant topic. the US public agenda also focused on this topic although not classifying as a true local minimum. In 2017, the US company Amazon acquired the US retailer *WholeFoods*, which is known for its large organic assortment.

In conclusion, agenda diversity differed significantly across the agendas. We have identified seven time periods of exceptionally low agenda diversity, which could be traced back to real world events. It is noteworthy that only *a* and *g* were minima in the

media agenda, whereas the remaining minima were in the public agendas. This confirms that the media agendas maintained a certain topic diversity in their reporting on organic food. In contrast, commenters temporarily focused their attention to specific topics, which were linked to political and economic events at the national and global level.

3.3.4.5 Agenda distance

In the previous section, we inspected the diversity of each of the media and public agendas separately. Agenda diversity was found to differ over time, with specific time periods exhibiting dominant topics. However, the topics did not always dominate media and public agendas to the same extent. Consequently, the question remains how similar those agendas were. We used the Jensen-Shannon distance (JSD) to measure the similarity between the relative distribution of topics in the media and public agenda for each country. Similar to the correlational metrics, JSD measured the relationship between articles and comments. However, while the correlational measures compared topic proportions individually, JSD considers the distribution of all topics. If JSD is close to 0, the topic distributions of articles and comments in this time period are very similar, i.e., the closer media and public agendas were related. At large values of JSD, the media and public agendas diverged. Results are displayed in Figure 10.

The insights from the agenda distance metric were manifold. First, we found rather low JSD values, indicating that the topic distribution was rather similar in articles and comments (see Figure 10). This confirms the findings from topic correlations above. Second, we detected differences in the agenda distance in both countries. The agenda distance was significantly higher in Germany compared to the US ($F(1, 143)=15.0, p < 0.001$). This suggests that German commenters stuck slightly less to the topics discussed in the media articles. Third, two local maxima were identified: *c* and *d* indicate periods of time where the public paid attention to different topics than the media 6. The time periods *c* and *d* coincide with the time periods of low agenda diversity in the public discussion in Germany (see Figure 8 and the topic distributions depicted in Figure 6). Agenda distance was large because—in contrast to the media—commenters had focused their attention on environmental policies in *c*, and animal welfare and economic issues in *d*.

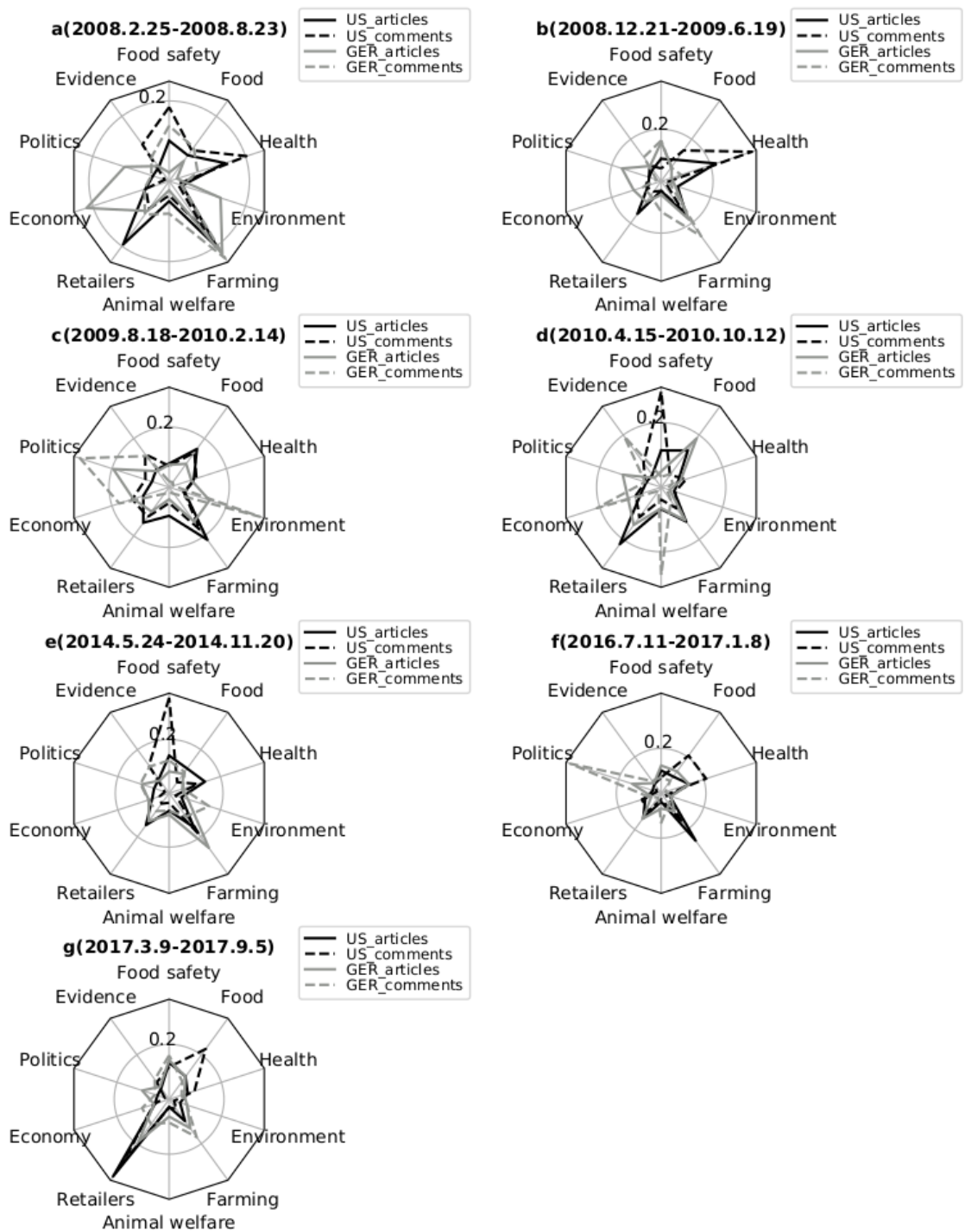


Figure 9 Topic distributions. Radar plots for the relative topic distribution in time periods *a, b, c, d, e, f,* and *g* identified via normalized Shannon entropy and JSD.

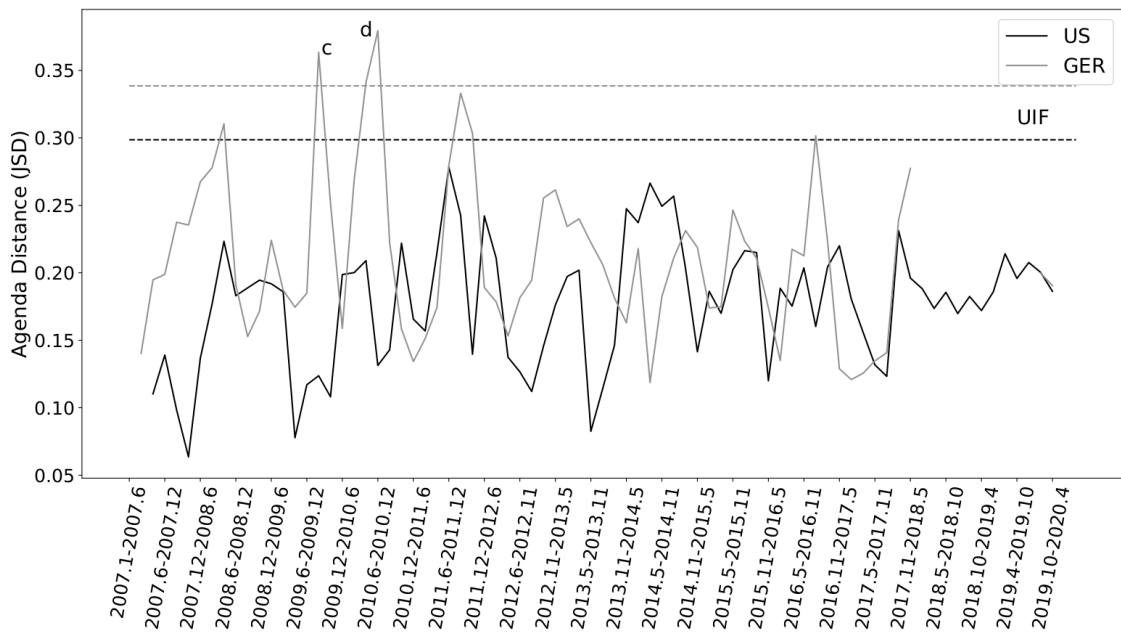


Figure 10 Agenda distance. Distance between media and public agenda for the US and Germany over time measured by Jensen-Shannon Distance. The horizontal lines point out upper inner fences (UIF) to identify local maxima, i.e., time periods with high agenda distance. These time periods are denoted with *c* and *d*, and coincide with the time periods depicted in the graph and radar plots of Figure 9. The horizontal lines point out the lower inner fences (LIF) to identify local minima, i.e., time periods with low agenda diversity.

3.3.5 Discussion

As expected, we found attribute agenda-setting relationships between the two news outlets and their audience. Both media outlets strongly influence the topics their readers write about: The topic proportions in news articles and comments were strongly correlated. Cross-lagged correlations revealed that the media agenda predicts the public agenda but not vice versa. While correlations—a classical agenda-setting metric—consider the proportions of each topic separately, we also accounted for the dynamics in the distribution of all topics on the agenda by calculating normalized Shannon entropy for agenda diversity and Jensen-Shannon distance for the distance between the media and public agendas. We showed that these metrics from information and probability theory provide additional insights into the dynamics within and between agendas (Boydston et al., 2014; Pinto et al., 2019). While confirming the close semantic relationship between articles and comments, they additionally pointed out a number of time periods where agenda distance was particularly high and where agenda diversity was especially low, because specific events caused the audience to focus on topics other than the media agenda. Moreover, we found that both agenda distance and agenda diversity were

significantly higher in Germany compared to the US. Further, in both countries, the media maintained a higher diversity than the public agenda.

Conclusions from this research are naturally limited to the two media outlets selected and their readership. In recent years, the growing number of online media outlets and social media has limited the agenda-setting power of single media outlets at a societal level and has led to a fragmentation of the public agenda (Takeshita, 2006). There is a multitude of different media and public opinion platforms, and agenda flows between them are increasingly reciprocal and dynamic (Denham, 2010; Neuman et al., 2014). Nevertheless, we found the influence between articles and comments to be unidirectional, since articles were not affected by prior reader comments. Research on how audience feedback is incorporated in news production is still scarce (Lee & Tandoc, 2017). However, reader comments matter: Cross-lagged correlations indicated that comments do influence future commenters. Readers potentially take the opinions stated in comments for the public opinion and conform to it (Lee, 2012). In a more recent study, Lee et al. (2017) showed that by making certain aspects of a news story salient, reader comments influenced how the event reported in the news is interpreted. Prior research has also shown that the tone of the comments influences subsequent comments (Ziegele et al., 2014) as well as the perception of the news article (Winter et al., 2015). Social network Twitter has been found to influence news media (Conway-Silva et al., 2018; Groshek & Groshek, 2014), although the traditional agenda-setting direction from the media to the public was still stronger. Future research could investigate the interplay between the news media and Twitter in the context of organic food. Meza and Park (2016) discovered that Twitter is a valuable channel for word-of-mouth communication on organic food both among consumers as well as between consumers and companies.

In addition to agenda-setting dynamics, our study provided a comprehensive overview of the landscape of topics salient in the media and their audience in the US and Germany. Topics ranged from typical product attributes of organic food such as food safety to production characteristics such as animal welfare as well as political and economic aspects to the organic food system. The validity of the identified topics is confirmed by previous results on topics salient to online users in the context of organic food (Danner & Menapace, 2020) as well as buying motives documented in consumer research (Hemmerling et al., 2015; Hughner et al., 2007; Testa et al., 2019). The ten topics

identified suggest that the media agendas in both countries have become more diverse compared to previous findings, which is likely to be linked to the growth of the organic food markets. The discussion in the US media from 1996 to 2002 had been limited to mainly food safety issues and the general conflict between organic and conventional foods (Lockie, 2006). In contrast, the results of a more recent study of several US print newspapers resound with our findings for the US context: the most salient topics are health, food safety, farming practices, and companies of the organic sector (Abrams et al., 2010). For Germany, an analysis of news media in the 1980s and 1990s revealed that the media coverage on agriculture in general focused on trade and agricultural policy—subsidies in particular (Gudrun Mahlau, 1999). Environmental aspects of agriculture were largely and organic farming completely ignored. By contrast, our findings imply that the current media attention is more diverse and takes into account the consumer perspective by discussing topics such as food safety, product quality, and healthiness.

3.3.6 Conclusions

By investigating news articles and related reader comments, our study confirmed the agenda-setting power of the news media (Conway-Silva et al., 2018; Groshek & Groshek, 2014). We uncovered that this agenda-setting function not only applies to political but also consumption contexts. Results indicate that the media can foster and direct their readers' discourse on organic food. Via agenda setting, the media determine issue and attribute importance of organic food, which in turn potentially influences on behavioral outcomes. Thus, the media possibly sustain a certain consensus-building function and carries ethical responsibility in selecting the issues and topics on the agenda (McCombs, 2014). The results of this study are relevant to different players in promoting organic food purchases. Marketers, lobbyists, policy makers, and politicians can consider the media as an important information channel by bringing specific topics to the media's attention, and thus shape the public perception of organic food. Media campaigns can accompany policy measures such as the introduction, adaptation, and promotion of organic standards, and thus increase the awareness, familiarity, and salience with consumers. Moreover, (Pollach, 2014) found that the news media coverage also determined the environmental agendas of companies. Future research could investigate whether this finding also holds true for the practices of the farming and food industry in particular. Undoubtedly, the

news media are an important player in the transition to a more sustainable food system, economy, and society.

3.4 Does online chatter matter for consumer behavior? A priming experiment on organic food (*Essay IV*)⁹

Abstract

Consumers are increasingly sharing their opinions on societal issues and products online. We studied the implications of such online word-of-mouth for consumer judgment and decision-making. The use case is organic food, which is the most successful among the currently emerging, sustainability-differentiated food product categories. First, we analyzed the online discussion on organic food by conducting a text-mining study of reader comments (N=63,379) from the comments section of a major German online news outlet. Topics therein are discussed with differing frequency, indicating the salience of the various issues to online readers and consumers. Based on the text-mining study, one organic food topic of high online salience (animal welfare) and one of low online salience (biodiversity) were selected for a priming experiment in a subsequent study. Second, we administrated an online survey to German consumers (N=1118). The included priming experiment investigated the behavioral relevance of online salient topics. In particular, we tested whether the relative online salience of the two selected topics used as primes influenced the likelihood of choosing organic instead of conventional eggs and milk in a choice experiment and the acceptance of policies supporting organic farming. While ineffective for milk choices, the priming worked as hypothesized with respect to policy acceptance and egg choices. Priming the topic with high online salience (animal welfare) is more effective at promoting pro-organic behavior than priming the topic with low online salience. Priming effects also depended on prime strength and were moderated by attitude strength and experience with buying organic food. Implications for policy and

⁹ A similar version of *Essay IV* has been published as Danner, H; Thøgersen, J. (2021): Does online chatter matter for consumer behaviour? A priming experiment on organic food. *International Journal of Consumer Studies*, DOI: 10.1111/ijcs.12732

CRedit Authorship Contribution Statement: **Hannah Danner**: Conceptualization, Methodology, Data curation, Validation, Formal analysis, Investigation, Resources, Writing - original draft, Writing - review & editing, Visualization, Project administration, Funding acquisition. **John Thøgersen**: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Writing - original draft, Writing - review & editing, Visualization, Supervision.

research are discussed, including: how insights from text mining of online word-of-mouth can be employed to promote sustainable consumption behavior.

Keywords: Priming; associative network; salience; text mining; user-generated content; consumer choices; organic food.

3.4.1 Introduction

People are increasingly sharing their opinions online. Today, 54% of the global populations have access to the Internet (ITU, 2019) and 51% use social media such as blogs, forums, and social networks like Facebook or Twitter (DataReportal, 2020). Worldwide, users spent a daily average of 2.5 hours online in 2018, a figure expected to increase to over 3 hours by 2021 (Zenith, 2019). The rise of online participation platforms has provided new sources of insight into consumer perceptions, attitudes, and behavior (Balducci & Marinova, 2018). In particular, text mining of online “chatter”, that is, online word-of-mouth or user-generated content (UGC; Tang et al., 2014; Zablocki et al., 2019), offers the possibility to “listen in” to discussions among consumers and find out what is on their minds. The widely available UGC online reveals the associations and product attributes that are salient to consumers (Gensler et al., 2015; Netzer et al., 2012). UGC is particularly valuable because it is created at the individual’s own initiative and without researcher influence, as opposed to, for example, interview data. The mere fact that users choose to express their concerns demonstrates the natural salience of these concerns and users’ involvement with the discussion topic (Ksiazek et al., 2016). Online comments are real expressions from real people aiming to express their views to other people on issues they are engaged in, which gives these expressions ecological validity. For the present purpose, we define “online salient topics” as associations mentioned in online discussions on a given issue. Whereas the volume of comments can be used as an alternative measure or indicator of the degree of user engagement (Ksiazek, 2018), the relative frequency with which a given topic is mentioned in online discussions can thus serve as an indicator of the relative salience of topics in the population online.

Sustainable consumption research has begun to analyze UGC online to investigate consumer perception. Researchers have taken stock of consumer beliefs and discussion topics and the relative salience of topics, including postings on sustainability-differentiated products like organic food (Danner & Menapace, 2020, Meza & Park,

2016; Olson, 2017). Research finds that the online sphere can influence consumer perceptions and sustainable consumption motives, for example, via (news) media consumption, peer-to-peer content and communication, and advertising (Cooper et al., 2012; Frick et al., 2020; Thøgersen, 2014). Frick et al. (2020) found that just “being” in online environments and the exposure to social media peer content and online advertisement influences sustainable consumption motives and behavior. Hence, with the current level of digitalization and Internet use, where also consumption patterns and their sustainability are increasingly discussed in and influenced by online environments (Chatzidakis & Mitussis, 2007; Reisch, 2001), research on sustainable consumption cannot ignore the online sphere.

However, there exists—to the best of our knowledge—little published evidence on whether *what* is discussed online and how much it matters beyond the online sphere, in particular, *how* the relative online salience of topics connects to everyday behavioral outcomes that are mostly offline, such as grocery shopping. This paper contributes to building knowledge in this unexplored area in several ways. First, it proposes a mechanism through which online salience might influence behavior from the perspective of priming and Spreading-Activation Theory. Second, it presents an empirical investigation of the relevance of the online salience of topics beyond the online sphere, as reflected in how consumers respond to being primed with a topic that is highly salient (vs. not very salient) online. According to Spreading-Activation Theory (Collins & Loftus, 1975; Fazio et al., 1986), priming increases the accessibility of associated knowledge, including evaluative knowledge, such as the attitude toward the product, and attitude activation impacts subsequent judgments and behaviors. With repeated exposure, the attitude becomes increasingly accessible and increasingly likely to influence behavior (Berger & Mitchell, 1989). We assume that a prime based on a topic that is very salient online activates related attitudes more than a prime based on a topic of low online salience. However, to be sure, we empirically test whether online salience is related to the accessibility of associative networks (i.e., whether topics that are more salient online have higher activation power) and thus matter for how consumers respond to emerging product categories.

Specifically, we test the effect of priming topics that vary in online salience on different behavioral outcome variables related to organic food. The investigated outcome variables

are the inclination to choose organic instead of conventional food in a choice experiment, and the willingness to accept organic farming policy measures. The priming is done by exposing participants in an online experiment to online news headlines. The topics included in the priming headlines were identified through a text-mining study, in which topics discussed in the online comment section of a major German newspaper were identified as having high versus low online salience, based on the number of user contributions mentioning the topic.

We use organic food as the case of study because organic food is the most successful representative of the emerging category of sustainability-differentiated products (Jackson et al., 2020; Willer et al., 2019), generating relatively many and engaged online expressions. Also, given that food consumption accounts for a large share of the environmental footprint of households (Poore & Nemecek, 2018; Tukker et al., 2010), understanding and promoting sustainable food consumption is of societal interest in its own right. Further, despite increasing availability, familiarity, and decreasing price premiums, sales of more sustainable foods have remained relatively low, suggesting a gap between positive attitudes and consumers' actual purchases (e.g., Aschemann-Witzel & Niebuhr Aagaard, 2014; Janssen, 2018; Schäufele & Hamm, 2018). Not least when it comes to food, it is necessary to explore reasons for the attitude-behavior gap and to find innovative approaches to promoting more sustainable purchases. In the food as in other areas, UGC might potentially provide novel insights into consumer rationales and into how to better promote emerging sustainability-differentiated food products (Danner & Menapace, 2020). In particular, there is a substantial online discussion on organic food and the search term "organic food" provides a feasible search term to identify and gather the relevant discussion data.

Next follows an overview of the theoretical basis of the study, including research objectives and hypotheses. Subsequently, we present a text-mining study that we use to identify the priming topics followed by the priming experiment itself. The paper ends with a general discussion and conclusion.

3.4.2 Theoretical background

It is now common to conceptualize the storage of consumer knowledge in long-term memory as associative networks, where concepts are saved as nodes, among which

information can spread (Bettman, 1979; Collins & Loftus, 1975). The spreading-activation model of associative memory suggests that external or internal cues that prime knowledge stored in memory increase the accessibility of stored knowledge associated with the prime (Janiszewski & Wyer, 2014; Minton et al., 2017). In a seminal paper, Fazio et al. (1986) showed that priming can automatically activate an attitude, which then affects subsequent product evaluations. Other studies based on this framework show that with repeated exposure to cues priming the same attitude, the attitude becomes increasingly accessible from memory and increasingly likely to influence subsequent behavior (Berger & Mitchell, 1989). Attitudes are conceptualized as object-evaluation associations in memory and as summary evaluations based on beliefs (i.e., appraisals of the attributes characterizing an object, as in an expectancy-value framework), affect, and/or behavioral information from past behaviors and experiences (Fazio, 2007). Exposure to an attribute object will automatically activate a respective attitude if the attitude is sufficiently accessible in memory. Fazio's (1986) Attitude-Activation Model suggests that attitudes guide behavior because such an automatically activated attitude serves as the starting point for what is emphasized in the immediate situation and, therefore, the evaluation of the attitude object, decisions, and behavior. Mounting research has confirmed that automatic evaluations can influence consumer judgment and decision-making (Petty et al., 2009). Such automatic evaluations are sometimes referred to as "implicit attitudes", which are usually inferred from the speed of participants' reactions to the combination of primes and evaluations (Petty et al., 2009). Fazio (2007) later summarized the process of automatic attitude-activation in the "model of attitudes as object–evaluation associations of varying strength".

In a further extension of the Attitude-Activation Model to a dual process model (the MODE Model), Fazio (1990) argued that behavior is guided solely by an automatically activated attitude whenever the motivation, ability or opportunity for further deliberation is lacking. However, when sufficient motivation, ability, and opportunity to deliberate are all present, a deliberate judgment entailing reasoned action is likely, as represented, for example, by the Theory of Planned Behavior (Ajzen, 1991; Ajzen & Fishbein, 2005). However, it is assumed that an automatic response is always activated first (cf., Kahneman, 2011).

Various priming measures have been shown effective at automatically activating attitudes (for an overview, see Fazio & Olson, 2003). For example, Shavitt and Fazio (1991) demonstrated that the attitude toward a product (i.e., an attitude object) can be automatically activated not only by exposure to the product itself, but also by exposure to an important product attribute. Priming specific attributes of an attitude object can affect subsequent behavior (Posavac et al., 1997; Shavitt & Fazio, 1991).

Information that is salient in one's environment can be "attention grabbing" and therefore can influence consumer judgments and decision-making (Tversky & Kahneman, 1974). By placing salient cues to relevant knowledge and attitudes in the environment, it is thus possible to "nudge" consumers toward a certain behavior (Reisch & Zhao, 2017; Thaler & Sunstein, 2009). Especially, personally relevant and vivid examples, explanations or cues are effective "attention grabbers" (Blumenthal-Barby & Burroughs, 2012). The effectiveness of cues in the environment also depends on how easy or difficult they are to process and how strongly associated they are to relevant knowledge and attitudes in receivers' memory (Keller, 1991), which is linked to frequency of prior exposure (i.e., familiarity) and direct experience (Berger & Mitchell, 1989).

In this paper, we use text mining to identify topics that are more or less salient in consumers' online environment, which is assumed to influence the frequency of prior exposure to a topic and therefore its effectiveness as a prime to activate relevant consumer attitudes, choices and behavior. Previous research has shown that the analysis of UGC contributes suggestive empirical evidence for which associated product attributes or topics related to the product are more and less salient to consumers. For the case of organic food, Danner and Menapace (2020) and Olson (2017) have analyzed the beliefs and topics discussed by users online showing that some topics are more salient in the online discussion than others.

This paper seeks to test whether these different levels of salience of topics in an online environment are relevant to consumer behavior offline. Based on Spreading-Activation Theory (Collins & Loftus, 1975) and Fazio's (2007) model of attitudes as object-evaluation associations of varying strength, we assume that topics that are very salient online have higher activation and spreading potential than topics of low salience because the associative networks that the former topics are part of, including relevant attitudes, have been more frequently activated in the past (e.g., by reading online news on organic

food and participating in the online discussion). Associations from large and interconnected associative networks have been shown to have high spreading potential (Nelson et al., 1993). The stronger the associations to a topic in memory, the more accessible are associated attitudes and the more likely is their automatic activation (Fazio et al., 1986; Fazio & Olson, 2003).

Against this background, we investigated whether the online salience of topics related to organic food influences attitude accessibility and, therefore, our ability to activate a favorable attitude and increase the likelihood of subsequent favorable behavior. In a priming experiment, we primed participants with only one of two topics related to organic food—either one that was very salient online, or one that was not very salient online. We investigated the effectiveness of the priming to activate a favorable attitude toward organic food as reflected in two, relevant behavioral outcome variables: choosing organic product alternatives (in a discrete choice experiment) and acceptance of policy supporting organic agriculture. The two outcome variables cover the perspectives of participants in their roles as consumers and citizens with regard to organic food-related behavior. Previous research has shown that participants are more likely to express their support for public goods like food safety or animal welfare when acting as voting citizens compared to buying consumers (Alphonse et al., 2014; Tonsor et al., 2009). Thus, considering voting behavior can take into account participants' support of the organic case even though they might not (frequently) buy organic products (e.g., due to the high price).

We expected that priming a topic that is strongly and positively associated with organic food is more likely to activate a favorable attitude toward organic food and therefore increase the likelihood that the person will take favorable actions toward organic food (cf., Fazio et al., 1986). Specifically, we hypothesized the following (see Figure 11):

H1: The choice of organic options and the acceptance of policy supporting organic food production can be increased by priming a topic that consumers associate favorably with organic food.

Building on this baseline hypothesis regarding priming, we further hypothesize regarding salience:

H2: The effectiveness of a priming intervention depends on two conditions: (a) the chronical salience of the priming cue or topic in the person's environment, including the

online environment (i.e., online salience), and (b) the situational salience of the priming cue or topic in the person’s immediate environment (i.e., prime strength).

According to Fazio’s attitude theory, the effects of a priming intervention will vary across individuals, depending on the strength of their attitude. The strength of the attitude matters and moderates information processing and evaluative judgments, with strong attitudes being more easily retrieved from memory (Fazio, 1995; Nayakankuppam et al., 2018). Fazio et al. (1986) argue that attitude strength determines attitude accessibility, that is, the ease with which an attitude is activated. Moreover, attitudes that are based on direct experience have been found to be stronger and more accessible than attitudes based on information alone (Berger & Mitchell, 1989; Fazio & Zanna, 1981), although repetition over time with regard to both direct experience and exposure to information implies more frequent activation and a richer cognitive network. Consistent with this, the frequency of past organic food consumption has been shown to positively impact buying intentions regarding organic food (Koklic et al., 2019). We therefore hypothesized the following:

H3: The effect of priming a topic that consumers associate favorably with organic food on the choice of organic options and the acceptance of a policy supporting organic food production depends positively on (a) attitude strength and (b) experience with buying organic food.

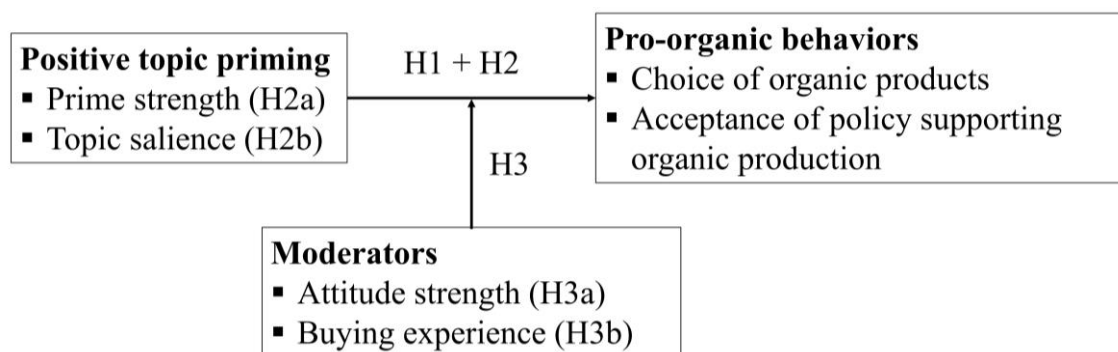


Figure 11 Research framework and hypotheses. Source: Own illustration.

3.4.3 Text-mining study on the online salience of organic food topics

We conducted a text-mining study to identify one organic food topic of high online salience and one of low online salience for the priming experiment (see Section 3.4.4).

3.4.3.1 Method

We analyzed 63,379 reader comments articles about organic food published online on *spiegel.de*, which is among the most visited news outlets in Germany (IVW, 2020). We used data from over a decade, from January 2007 to December 2017,¹⁰ to ensure that our measure of the relative salience level of the topics was not overly influenced by single events at any given point in time. We used the well-established Latent Dirichlet Allocation (LDA) algorithm (Blei, 2012) to model the topics. Topic modeling is a form of unsupervised machine-learning which discovers latent semantic structures (i.e., topics) and models the probability of words in documents (i.e., comments). Topics are clusters of semantically similar and co-occurring words. The topic model yields a weighted combination of the different topics that represent a document. We used the relative share of comments that made mention of a topic as our indicator of the relative salience of topics.

3.4.3.2 Results

Our text mining identified 18 different topics on organic food in the reader comments in *spiegel.de*. The relative salience of topics differed substantially with *politics* and *animal welfare* being the most salient topics (Figure 12). An analysis over time showed that the relative salience of topics varied, but, despite certain peaks, it was quite stable over the decade covered by the text analysis. The average salience and salience stability indicate that some topics related to organic food are perennially more salient in this online environment than others.

¹⁰ Unfortunately, comments on *spiegel.de* in 2018 and 2019 were not available due to a website restructuring.

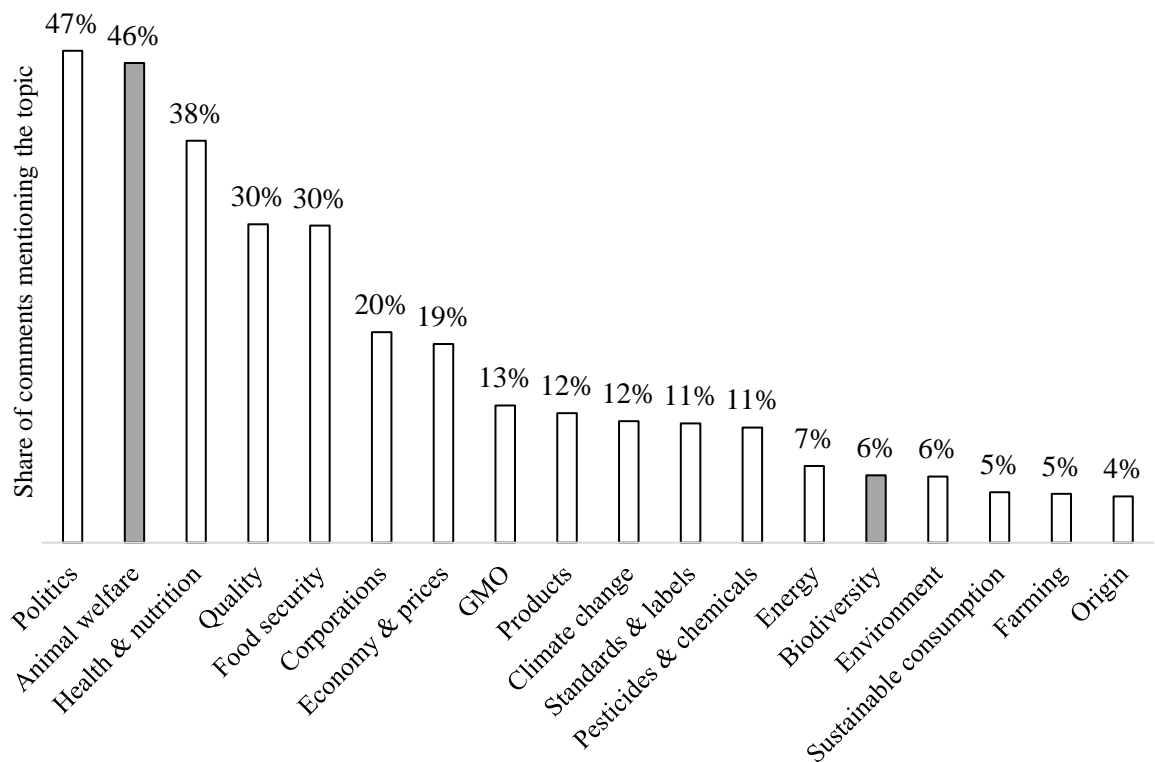


Figure 12 Overall share of comments on organic food mentioning each topic (2007-2017) in *spiegel.de*. The selected topics *animal welfare* and *biodiversity* are highlighted in grey. Source: Own illustration.

3.4.3.3 Selection of organic food topics for the priming experiment

Based on the text-mining study, we chose *animal welfare* as a relatively highly salient topic and *biodiversity* as a topic of low salience in relation to organic food. These topics were to be used as primes in the priming experiment (see Section 3.4.4). From Figure 13, we see that the two topics showed consistently different salience levels over time, with *animal welfare* consistently being far more salient than *biodiversity*. Overall, *animal welfare* was represented in 46% of all comments and *biodiversity* in only 6% of comments (Figure 12). This is in line with previous qualitative UGC analyses finding that *animal welfare* is more than twice as salient in online user comments on organic food than *biodiversity* (Danner & Menapace, 2020).

The following reasons assured the comparability of both topics and further supported their suitability for the priming experiment. (1) *Animal welfare* and *biodiversity* are both included in national and EU organic regulations (e.g., on animal husbandry, pesticide use, extensive land use, and crop rotation). (2) The respective benefits of organic over conventional farming have been documented (Spooler, 2007; Tuck et al., 2014), as

opposed to other aspects like healthiness, which is a main reason for buying organic food according to many studies (e.g., Rana & Paul, 2020) but evidence on actual health impact is ambiguous (Smith-Spangler et al., 2012). (3) In consumer perception of organic food, both of these topics generally have a positive connotation, (4) and they both refer to public rather than private benefits of organic food. (5) Both topics are highly important product attributes of organic food in the eyes of consumers. In a recent consumer survey in Germany (Federal Office for Agriculture and Food, 2020), animal welfare was rated as important in the context of organic food by 95% of participants, and biodiversity by 85% of participants.

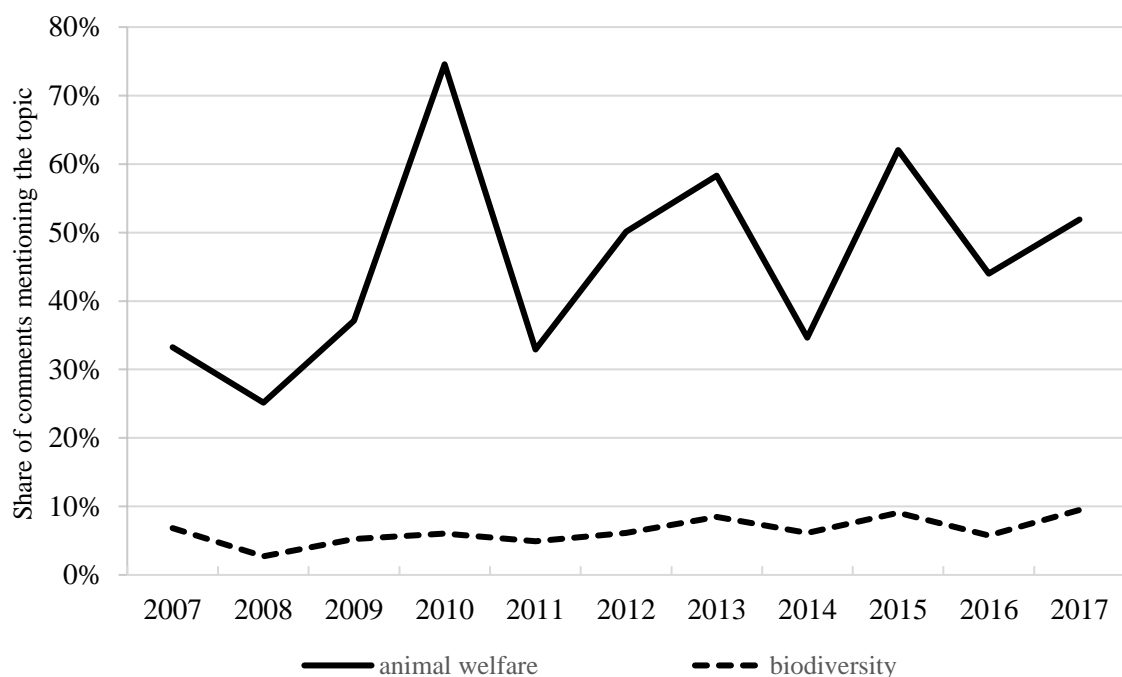


Figure 13 Share of comments on organic food mentioning the topics *animal welfare* and *biodiversity* over time (2007-2017) in *spiegel.de*. Source: Own illustration.

3.4.4 Priming experiment

Using an online survey, we analyzed effects of priming the two mentioned organic food-related topics with online salience according to the text-mining study (*animal welfare* and *biodiversity*). Specifically, we investigated how the priming of the two topics influenced the willingness to choose organic food when conventional options and the acceptance of a policy to promote organic food.

3.4.4.1 Methods

The data were collected in Germany by means of an online survey in October 2019. The questionnaire was programmed in Qualtrics and administered by a market research company. A quota-sample (N=1118) was drawn so as to be representative of the German adult population in terms of gender, age, schooling, professional education, and income. See Appendix 5 for the socio-demographic respondent characteristics. We screened participants to make sure they were at least partly responsible for grocery shopping in their households and that they bought the two products used in the priming experiment at least occasionally. The screening was meant to ensure that the investigated behavior was relevant to participants and thereby increase answer quality.

We randomly assigned participants to priming and control groups (see 4.1.1) and to two choice experiments (eggs and milk, see 4.1.3). Comparisons of means confirmed that there were no relevant differences between the priming and control groups or between the eggs and milk group¹¹ regarding elicited sample characteristics (socio-demographics, attitude valence and strength, buying experience with organic food, importance of the priming topics, media use, and trust; see Appendix 6 for an overview of all elicited variables). Prior to the survey, the questionnaire and priming mechanism had been pre-tested with participants recruited from the crowd-working platform Prolific (N=240).

3.4.4.1.1. *Priming*

We conducted a priming experiment using a between-subjects design where participants were randomly assigned to four different priming groups and a control group. We exposed participants in the priming groups to one of the topics determined in the text-mining study. The selected topics were *animal welfare* and *biodiversity* as outlined in Section 3.4.3.3. Participants exposed to each topic were further divided into two groups: a strong and a weak priming group, in order to investigate the importance of the situational salience of the priming or prime strength. This resulted in five groups in total: four priming groups and a control group. The priming was done by exposing participants to seven headlines. In the control group, the (“neutral”) headlines were about topics unrelated to farming,

¹¹ At a Bonferroni-corrected level, only 3 out of 14 sample characteristics variables were significantly different in the eggs and milk group and the effect sizes were small. To save space, analyses were not reported but can be acquired from the authors.

food, animal welfare, or biodiversity. In the weak/strong priming conditions, one/three of the seven neutral headlines were replaced by headlines related to either animal welfare (the topic of high online salience) or biodiversity (the topic of low online salience). See Appendix 7 for the priming scenario and items. The priming interventions were expected to activate the person's associative network regarding organic food and therefore increase the accessibility of, attention to, and preference for organic food (e.g., when choosing between organic and non-organic alternatives) to differing degrees, whereas the control group was supposed to reveal the baseline effect of participants' predispositions.

3.4.4.1.2. Procedure

After being welcomed and screened, the participants answered socio-demographic questions, followed by the priming. Then we asked whether they would vote for a policy that would redistribute agricultural subsidies according to criteria favorable to the environment, the climate, and animal welfare, thus supportive of organic agriculture (Alphonse et al., 2014; Tonsor et al., 2009). See Appendix 6 for the wording of the policy acceptance items as well as all other employed items and scales. Next, participants made 12 choices between products, either eggs or milk. Then, as filler questions, we elicited the frequency of nine different media consumption behaviors and trust in media outlets. Furthermore, we measured attitude valence (the overall liking of organic food) and attitude strength on a seven-item scale: importance, centrality, identity, representativeness, certainty, subjective knowledge, and likelihood of change (Pomerantz et al., 1995). Finally, we assessed past experience with buying organic food by asking about how long they have been buying organic food. The questionnaire concluded by eliciting the importance of the topics of animal welfare and biodiversity in the context of organic food.

3.4.4.1.3. Choice experiment design

We measured the impact of the priming intervention on consumer choices in discrete choice experiments in which randomly assigned participants made choices between alternative options of either egg or milk products (to learn whether found effects are product-specific). We framed the choice situation as follows: "Imagine you are grocery shopping. You have the choice between different alternatives of milk/eggs. They differ regarding their packaging, their price, their origin, and whether they were organically or

conventionally produced.” The choice experiment consisted of 12 repeated choices between two options for 10 eggs or 1 liter of milk. Participants were also free to choose “none of these” if none of the options were appealing. The products varied in terms of: *organic* versus non-organic, *price* (eggs: €1.49, €2.39, €3.29, €4.19; milk: €0.69, €0.99, €1.29, €1.59, anchored by market prices at German retailers), *origin* (local, Germany, EU), and *packaging* (milk: normal Tetra Pak, Tetra Pak made from recycled materials; eggs: tray, loose). We used *Ngene* software to create a d-efficient choice design (Street et al., 2005), resulting in 24 choice sets that were blocked into two equally large blocks of 12 choice sets each (see Figure 14 for examples), to which participants were randomly assigned in order to limit cognitive effort and prevent fatigue (Swait & Adamowicz, 2001).

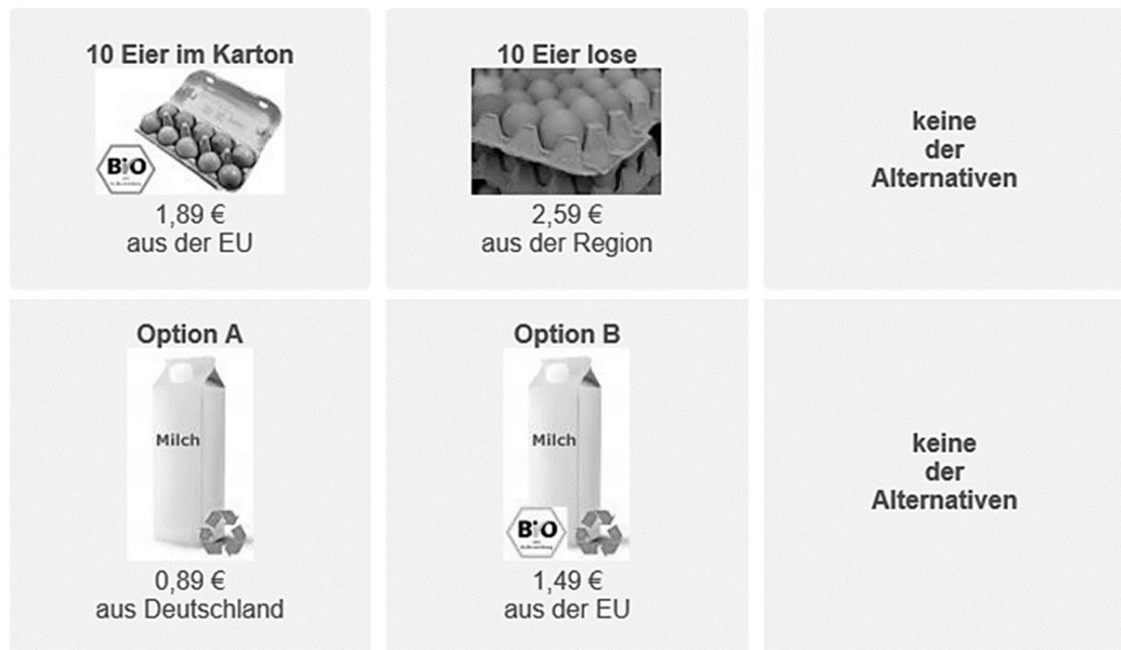


Figure 14 Examples of the eggs and milk choice sets. Source: Own illustration.

3.4.4.2 Results

The following sections report the results of the priming experiment for each of the analyzed outcome variables, that is, choices and policy acceptance moderated by attitude strength and buying experience.¹²

¹² Moderation by attitude valence was also tested and effects pointed in the same direction as attitude strength and buying experience.

3.4.4.2.1. *Choice experiment*

We estimated an extended multinomial logit model using *Latent Gold 5.1* (Vermunt & Magidson, 2016), including the attributes of the alternatives as independent variables (McFadden, 1974), separately for eggs and milk. Some participants in choice experiments make random choices without considering the alternatives, which can lead to biased parameter estimates (Grunert et al., 2015; Thøgersen et al., 2019; Thøgersen & Alfinito, 2020). In *Latent Gold*, it is possible to identify such random choosers by defining a latent class for which all levels of all attributes have equal utility, that is, none of the attributes matter for the choices made. Across the estimated models, approximately a fifth of the participants turned out to be random choosers, which is a considerable but not an exceptional share (Grunert et al., 2015). When testing the effects of priming conditions, the four priming groups and the control group are defined as known classes. We tested (1) whether the importance of attributes and attribute levels differs between priming groups and the control group, and (2) whether estimated relationships are moderated by attitude strength and buying experience. *Price* is defined as a continuous variable, and all other attributes as nominal. The reference levels are EU for *origin*, non-organic for *organic*, and non-recycled Tetra pack/tray for *packaging*.

Four models are presented in Table 10—two for eggs and two for milk, one of each including attitude strength and the other the duration of participants' experience with buying organic food as moderator. For all of the models presented, the Wald statistics show that choices were significantly influenced by *organic*, *packaging*, *origin*, and *price*. As expected, lower prices were preferred over higher, organic over non-organic, more sustainable packaging over less, and local origin over national, which was preferred over EU origin. Since the priming intervention was specifically meant to influence the importance given to *organic*, the interpretation of results focuses on this attribute in particular. The Wald (=) statistic—testing the equality of part worth utilities (i.e., regression parameter estimates) across groups—shows that they differ significantly for *organic* across the five conditions. A supplementary paired Wald test is used to identify which specific conditions led to significantly different utilities for *organic*.

For the choice of eggs (Models 1-2), we find that the online salience of topics used as primes matters for the effectiveness the priming. As hypothesized (H1), in both Model 1 and 2, receiving a strong *animal welfare* priming (the topic with high online salience)

significantly increased the likelihood of choosing *organic*. The strong *animal welfare* priming was the only condition in which the preference for *organic* was significantly different from and higher than in the control group. Furthermore, in this condition, the preference for *organic* was also significantly higher than in the two *biodiversity* priming conditions. Hence, it appears that both a high online salience of the topic used for the priming (i.e., *animal welfare*, H2a) and a high situational salience (i.e., strong priming, H2b) are necessary for significantly increasing the choice of organic eggs, only partly confirming H2a and H2b. The employed weak prime was not sufficient, even when the topic had high online salience. Priming the low-salience online topic of *biodiversity* was not effective at increasing organic choices, even with a strong priming.

For the choice of milk (Models 3-4), the priming effects were not as hypothesized. Inconsistent with H1, the choice of *organic* was highest in the control group, suggesting no priming effects. In Model 3, the priming effects of weak *animal welfare* and strong *biodiversity* were significant but negative. The other priming conditions did not differ significantly from the control group. In Model 4, the choice of *organic* in all priming conditions was significantly smaller than in the control group. Therefore, it seems that the effects of priming online salient topics on organic choices differ between products, specifically between eggs and milk.

Consistent with H3, we found—for both the egg and milk choices—that having a strong attitude toward organic food (H3a) and longer buying experience (H3b) increased the likelihood of choosing *organic* across all conditions. The effect of the moderators was strongest when *animal welfare* was strongly primed, suggesting that the effect of priming an online salient topic is highest among consumers with strong attitudes about the product in question and/or substantial buying experience.

Table 10 Latent class multinomial logit analyses for the choices of eggs and milk for the four priming conditions (weak/strong animal welfare (AW), weak/strong biodiversity (BD)) and the control group (CG) moderated by attitude strength and experience with buying organic food.

Prime topic Prime strength	Animal welfare		Biodiversity		Control	Wald (0)	Wald (=)	Wald (paired)
	weak	strong	weak	strong				
Model 1: Egg choices moderated by attitude strength (N = 633; 22% of random choosers excluded; R ² = 0.321)								
Class size	95	97	107	95	101			
<i>Intercept</i>	-0.083	-0.058	0.010	-0.083	-0.013	5.580 ^{ns}		
<i>Organic</i>	0.488	0.535	0.394	0.328	0.386	429.720***	12.741*	weak AW>strong BD* strong AW>weak BD* strong AW>strong BD** strong AW>CG*
<i>Packaging</i>	0.052	0.149	0.089	0.078	0.182	39.880***	6.027 ^{ns}	
<i>Germany</i>	0.212	0.117	0.111	0.345	0.124	1150.536***	16.407*	
<i>Local</i>	0.789	0.986	0.911	0.706	0.984			
<i>Price</i>	-0.874	-0.642	-0.757	-1.139	-0.629	737.873***	36.106***	
<i>None</i>	-2.410	-2.707	-3.501	-3.347	-2.534	884.534***	594.856***	
<i>Organic</i> <i>*attitude strength</i>	0.279	0.443	0.430	0.269	0.308	243.770***	10.305*	weak AW<strong AW* weak AW<weak BD* strong AW>strong BD* weak BD>strong BD*
Model 2: Egg choices moderated by buying experience (N = 633; 24% of random choosers excluded; R ² = 0.300)								
Class size	90	91	109	96	94			
<i>Intercept</i>	-0.136	-0.137	0.061	-0.092	-0.086	20.770***		
<i>Organic</i>	0.389	0.640	0.436	0.377	0.402	407.625***	12.892*	weak AW<strong AW** strong AW>weak BD** strong AW>strong BD** strong AW>CG**
<i>Packaging</i>	0.090	0.168	0.053	0.030	0.182	34.460***	8.781 ^{ns}	
<i>Germany</i>	0.164	0.105	0.160	0.329	0.180	938.052***	33.937***	
<i>Local</i>	0.845	1.153	0.855	0.580	1.011			
<i>Price</i>	-0.985	-0.481	-0.675	-1.097	-0.684	863.569***	63.397***	
<i>None</i>	-2.906	-2.565	-3.319	-3.077	-2.655	760.956***	517.012***	
<i>Organic</i> <i>*buying experience</i>	0.293	0.321	0.241	0.028	0.233	99.703***	14.118**	weak AW>strong BD** strong AW>strong BD*** weak BD>strong BD** strong BD>CG*
Model 3: Milk choices moderated by attitude strength (N = 485; 21% of random choosers excluded; R ² = 0.350)								
Class size	74	68	73	79	89			
<i>Intercept</i>	-0.062	-0.153	-0.075	-0.011	0.110	5.176 ^{ns}		
<i>Organic</i>	0.542	0.725	0.644	0.574	0.746	685.382***	11.591*	weak AW<strong AW* weak AW<CG** strong BD<CG*
<i>Packaging</i>	0.357	0.340	0.269	0.282	0.292	182.518***	2.158 ^{ns}	
<i>Germany</i>	0.096	0.216	0.305	0.091	0.121	1025.156***	8.509 ^{ns}	
<i>Local</i>	1.051	0.909	1.002	1.077	0.976			
<i>Price</i>	-0.506	-0.763	-0.781	-0.526	-0.599	598.472***	17.698**	
<i>None</i>	-3.607	-2.872	-4.057	-4.592	-2.683	361.270***	235.596***	
<i>Organic</i> <i>*attitude strength</i>	0.180	0.529	0.375	0.438	0.446	192.238***	19.938***	weak AW<strong AW** weak AW<weak BD* weak AW<strong BD*** weak AW<CG***

Prime topic Prime strength	Animal welfare		Biodiversity		Control	Wald (0)	Wald (=)	Wald (paired)
	weak	strong	weak	strong				
Model 4: Milk choices moderated by buying experience (N = 485; 15% of random choosers excluded; R2 = 0.305)								
Class size	79	78	88	80	86			
<i>Intercept</i>	-0.030	-0.030	0.090	-0.010	0.069	1.453 ^{ns}		
<i>Organic</i>	0.410	0.555	0.495	0.568	0.760	608.488***	22.281***	weak AW<strong AW* weak AW<strong BD* weak AW<CG*** strong AW<CG** weak BD<CG*** strong BD<CG*
<i>Packaging</i>	0.332	0.243	0.247	0.248	0.241	171.110***	2.876 ^{ns}	
<i>Germany</i>	0.075	0.199	0.211	0.120	0.160	1018.013***	11.375 ^{ns}	
<i>Local</i>	0.852	0.771	0.890	1.011	0.944			
<i>Price</i>	-0.499	-0.559	-0.558	-0.430	-0.510	614.971***	6.243 ^{ns}	
<i>None</i>	-3.902	-2.856	-3.303	-5.459	-2.900	457.985***	366.565***	
<i>Organic</i> <i>*buying experience</i>	0.113	0.361	0.066	0.285	0.200	123.172***	32.179***	weak AW<strong AW*** weak AW<strong BD* strong AW>weak BD*** strong AW>CG* weak BD<strong BD*** weak BD<CG*

Source: Own table.

Notes: Parameter values for random choosers are by definition fixed to zero, so they were omitted.

For the paired Wald tests, we only reported significant pairwise comparisons for the *organic* attribute and the interaction term. For the paired Wald tests, “animal welfare” is abbreviated with AW, “biodiversity” with BD, and “control group” with CG. * p < 0.05, ** p < 0.01, *** p < 0.001

3.4.4.2.2. Policy acceptance

In addition to the choice experiment, we test the effect of the *animal welfare* and *biodiversity* primes on participants’ acceptance of a policy regarding organic food. We estimate an ANCOVA in *SPSS 26* with the priming conditions as between-subjects variable and attitude strength and buying experience as covariates and moderators. We also include additional potential predictor variables as covariates (the perceived importance of animal welfare and biodiversity, and media trust). The results are shown in Table 11.

Inconsistent with H1, the main effects of the different kinds of priming on policy acceptance are not significant ($F(4, 1113)=1.452, p=0.215$; see Table 11). However, consistent with H3a, attitude strength significantly moderates the effect of the priming ($F(4,1113)= 2.379, p=0.050$). Planned contrasts revealed that, among the participants with the strongest attitudes toward organic food, all primes except the weak *animal welfare* prime ($t(1106)=1.659, p=0.081$) significantly increased policy acceptance compared to the control group. Consistent with H2a, the strong *animal welfare* priming showed the largest priming effect (strong *animal welfare* prime, $t(1106)=2.689, p=0.007$;

weak *biodiversity* prime, $t(1106)=2.454$, $p=0.025$; strong *biodiversity* prime, $t(1106)=2.216$, $p=0.045$). Contrary to expectations (H3b), the interaction between the priming conditions and buying experience was not significant.

Table 11 ANCOVA for policy acceptance moderated by attitude strength (N = 1118; R2 = 0.279, Ref = control group).

	B	SE	t	p
Weak animal welfare	-0.122	0.075	-1.618	0.105
Strong animal welfare	-0.168	0.071	-2.256	0.024
Weak biodiversity	-0.102	0.076	-1.369	0.174
Strong biodiversity	-0.135	0.074	-1.793	0.065
Weak animal welfare * attitude strength	0.122	0.070	1.659	0.081
Strong animal welfare * attitude strength	0.205	0.069	2.689	0.007
Weak biodiversity * attitude strength	0.180	0.076	2.454	0.025
Strong biodiversity * attitude strength	0.166	0.082	2.216	0.045

Source: Own table.

Notes: The significance of effects and standard errors (SE) was determined through bootstrapping with 1000 bootstrap samples.

3.4.5 Discussion

The results of this study confirm that reminding consumers of topics (i.e., priming) that are salient online can be a useful tool to boost product choices (in this case, of organic eggs) and policy acceptance (in this case, policy supporting organic agriculture). As expected, priming a topic that is very salient online—*animal welfare*—is more effective than priming a topic that is less salient—*biodiversity*. This is consistent with the assumption that the priming of online salient topics is more effective at activating or activates a stronger associative network in consumers' minds, therefore more strongly influencing their judgments and behavior. Thus, the online salience of topics seems to be an important aspect of the “choice architecture” for consumer choices and the priming of salient topics can be an effective “nudge”. Besides the online salience of the prime, the effectiveness of priming appears to depend on the situational salience of the priming, or prime strength. A stronger prime appears to be more effective than a weaker, suggesting that—to be effective—the priming should not be too subtle. This is in line with the findings of a recent review of salience and priming interventions to nudge people toward healthier choices (i.e., providing relevant health information) (Wilson et al., 2016). This review found that salience nudges alone show mixed effectiveness, but they become more effective when combined with other nudges and interventions, such as modifying the visibility, availability, and/or accessibility of choices.

Furthermore, the effect of priming on organic food choices and organic policy acceptance appears to depend on the strength of attitudes toward organic food and the effect on choices also on experience with buying organic food. This is consistent with the assumption that the strength of the primed associative network matters for the susceptibility to primes and thus the likelihood of attitude activation.

The results confirmed our hypotheses about organic policy acceptance and also with regard to product choices when the product was organic eggs. However, when the product was organic milk, the impact of priming on choices was contrary to expectations. The different results for eggs and milk are surprising since both products are among the most-purchased organic products in Germany.¹³ In addition, both eggs and milk are often among the first products that consumers choose when they start buying organic products (Juhl et al., 2017). The selected priming topics, *animal welfare* and *biodiversity*, are both important reasons for why consumers purchase organic food in Germany (Federal Office for Agriculture and Food, 2020). We would therefore assume the motives for buying organic eggs and organic milk to be similar. For example, using US scanner data, Lusk (2011) did not find substantial differences in how food values explained the demand for organic eggs and milk in the USA. Studies have also found that German consumers are willing to pay the same price premium for eggs and milk produced using enhanced animal welfare standards (Heise & Theuvsen, 2017).

We can only speculate why results differ between the two products. Priming of specific topics comes with uncertainty in the selection of these topics. First, behaviors can only be expected to be consistent with the prime if consumers deem the prime relevant to the subsequent behavior or judgment (Janiszewski & Wyer, 2014). It is possible—but nevertheless surprising since milk and eggs are rather similar organic products—that participants did not perceive the two primed topics as being relevant to their choice of organic milk (DeCoster & Claypool, 2004). So, even if they cared about animal welfare and biodiversity as such, they might not have believed that buying organic milk was relevant to these issues. Second, the higher inclination to choose organic milk in the control group than in the priming groups might indicate that the attributes spontaneously

¹³ Organic eggs and milk are among the top three organic products in Germany in terms of sales volume, amounting to 14.1% of total egg sales and 10.1% of milk sales (Organic Food Production Alliance [BÖLW], 2020).

associated with organic milk were more influential than the associations produced by priming *animal welfare* or *biodiversity*. Participants could have had many personal cues and a large associative network for organic food because the German market for organic milk is rather mature. Milk could also have been more a habitual choice in comparison to eggs, with eggs being less frequently purchased than milk,¹⁴ with consumers thus paying more attention to the attributes of the products offered in the case of eggs. *Animal welfare* and *biodiversity* could also have been already very accessible in participants' minds when choosing milk and therefore the priming makes no difference. Furthermore, personal associative networks could have been dominated by more self-interested and potentially more meaningful motives (e.g., price, healthiness), as opposed to the altruistic topics of *animal welfare* and *biodiversity*. Some studies have found that priming personal gain in contrast to environmental gain can encourage sustainable consumption behavior (Chwialkowska & Flicinska-Turkiewicz, 2020). Third—and in contrast to the previous arguments—there could be some kind of backfiring or boomerang effect regarding the choice of milk. According to Psychological Reactance Theory, persuasion attempts can be perceived as a threat to personal freedom and can lead to the opposite of the desired behavior (Brehm, 1966). If raising the awareness for *animal welfare* and *biodiversity* issues in food production is perceived as pressure, this could have caused some individuals to choose fewer organic options. Furthermore, Attribution Theory suggests that individuals justify not having behaved as desired because they perceive the issue as being outside of their control (Kelley, 1967). People might then resort to a type of “task difficulty excuse” (Bira & Faure, 2018). However, it is not clear why reactance should be produced when making choices involving organic milk products, but apparently not when making choices involving organic eggs. We will leave it for future research to solve this puzzle.

3.4.6 Conclusion and implications

The present study contributes insights on how understanding what is salient in online environments, and text mining of UGC online, can be employed to promote sustainable consumption behavior, such as buying organic food. It produced new insights regarding

¹⁴ An analysis of German household scanner data from Gesellschaft für Konsumforschung (GfK) shows a higher monthly frequency of milk purchases (M=3.01, SD=2.44; 2016-2017) than that of egg purchases (M=2.12, SD=1.33; 2012-2014).

the link between the online salience of topics and behavioral judgments, decision making, and behavior. However, further research is needed on the benefits of priming online salient topics in communications promoting products or causes. Although it appears that both the online salience of topics and the strength of the priming matter, our results are somewhat ambiguous. Therefore, future research should analyze more products and product categories to further investigate possible product or category-specific differences in the effectiveness of priming based on online salience.

Nevertheless, our results support the common perception that what goes on in online environments, including the salience of topics and online word-of-mouth, actually impacts consumer judgment and choices. When explicitly elicited, German consumers rate the importance of both of the included attributes (*animal welfare* and *biodiversity*) as high. In a recent consumer survey, 95%/85% of participants found *animal welfare/biodiversity* to be important in the context of organic food (Federal Office for Agriculture and Food, 2020). Hence, survey research suggests a far smaller difference in importance between the two topics compared to the difference in natural salience as attested by our text-mining study, where *animal welfare* was found to be far more salient to users than *biodiversity*. Thus, despite the mixed effectiveness of the priming for the choice of organic in different product categories, our results suggest that online salience matters. Priming with a more salient topic caused more pro-organic behavior than priming with a less salient topic. The detection of the power of online “chatter” to activate product-related knowledge and attitudes means that we should listen more to people’s online expressions in order to find out what are salient issues at present, and then use this insight to promote desired choices and behavior. This is relevant not only with regard to products but also with regard to acceptance of policy to promote desired behavior. Future research could investigate this relationship for other products, product categories, different sources of UGC online, and, not least, its applicability to other cultural contexts.

The findings entail important practical implications for marketing and policy communication, and public opinion polling. The revealed impact of priming online salient topics and voter acceptance of organic agriculture policy suggests that it is worthwhile to monitor online discussions. What is salient online can be an important indicator of popular political opinions on contemplated policy and regulation. For example, policy makers involved in the introduction, adaptation, and promotion of sustainability

standards, can use insights on online discussions to guide communication strategies aiming to increase the awareness and acceptance of these issues with citizens.

Marketing communication can also benefit from monitoring online word-of-mouth about relevant product attributes and product positioning. Communication could synchronize messages with what is salient in relevant online fora to increase their effectiveness, building on the insight that salient topics are more accessible and can be primed more easily. Communicators might also try to make specific topics more salient. The present text-mining and priming studies have demonstrated that *biodiversity* is currently not very salient in German consumers' online environment, so priming this topic does not influence consumer choices and behavior a great deal. However, *biodiversity* is critical to the resilience of our earth system (Steffen et al., 2015) and it is a core benefit of organic farming, which has shown clear advantages over conventional farming in this respect (Tuck et al., 2014). As a result, marketers and policy makers might highlight more such verifiable benefits of organic production and mobilize the agenda-setting power of the media to do so (Thøgersen, 2006). News media can be an effective channel for marketers and policy makers to reach consumers and bring core aspects into the public debate and onto their agenda, and potentially influence sustainable consumption behavior (Bellotti & Panzone, 2016).

4 Conclusion

This section provides a conclusion to the research of this dissertation. First, Section 4.1 will summarize and discuss the main findings. Subsequently, implications for future research as well as managerial and policy implications will be presented in Sections 4.2 and 4.3.

4.1 Main findings and discussion

Current consumption patterns critically challenge the environment. In response, sustainability-differentiated product categories are emerging. In the context of food production—an important contributor to climate change—, organic food is currently the most successful sustainability-differentiated product category. Certification systems have been established, organic products are widely available, consumer prices have decreased, and market shares are rising globally. However, understanding respective consumer perception and behavior is key to further increase organic purchases and tap the sales and sustainability potential. At the same time, digitization enables new platforms for consumer discourse and makes consumer reasoning visible online: Users discuss their concerns, reveal their consumption motives, and evaluate products on online platforms. Such UGC has become an abundantly available data source, which—as shown in this thesis—can be exploited by consumer research to understand consumer perception and behavior.

Contribution. This dissertation investigated consumer perception and behavior with regard to sustainable consumption using the example of organic food. It examined organic food beliefs and topics salient on consumers' minds by analyzing different types of UGC using various qualitative and quantitative text mining methods. In four consecutive essays, the thesis addressed the following four research questions:

Essay I: Which beliefs and topics about organic food are salient to online commenters in German-speaking countries and the US?

Essay II: How can qualitative content analysis and machine-learning-based analysis methods be combined to yield validation?

Essay III: Are there agenda-setting effects between the media and their audience regarding salient organic food topics?

Essay IV: Can topics salient online be used to prime consumer judgment and decision-making regarding organic food?

This dissertation contributes to consumer research, in particular research into sustainable and pro-environmental behavior. It stands out due to its interdisciplinary, mixed-method approach with respect to theories, data, and methods. To gain insights into online consumer perception of sustainable consumption and its link to consumer behavior, this thesis considered different theoretical angles, ranging from Spreading-Activation Theory, to consumer preferences, and Agenda-Setting Theory.

This thesis triangulated different types of data and methods across the four essays to yield valid and interpretable findings for consumer research considering different perspectives (Denzin, 1978). The text-mining studies of this thesis used both small and large data sets of UGC and applied both established consumer research methods such as content analysis as well as novel big data approaches based on machine-learning and neural networks. Finally, survey research is conducted including a choice experiment. This research analyzes data from the German-speaking countries and the US, the largest organic markets globally. In addition to data from Germany, *Essay I* also considers data from Austria and Switzerland. *Essays I* and *III* involved a comparison between the two language contexts, whereas *Essay II* focused on US and *Essay IV* on German data. The dissertation considered different players: Online users and consumers voicing their opinions online were the main target as all four essays comprised text mining of textual UGC. In addition to a preliminary text-mining study, *Essay IV* took into account the general public by testing the relevance of online salient topics for a sample representative of the German adult population. *Essay III* took into consideration the news media as important players and agenda setters.

The text-mining analyses revealed that consumers discuss organic food vividly representing a diversity of positions and angles. Illustrated by the example of organic food, this finding testifies their arousal and concern with pressing societal issues related

to sustainable food production and consumption. In sum, this dissertation contributes to the understanding of current consumer perception of organic food in terms of beliefs and topics salient online, the influence of the media on the public agenda, and the use of online salience to determine pro-organic consumer behavior. The findings of the four essays will be summarized in the following.

Essay I. The first essay presented an inventory of consumer beliefs about organic food in online comments from German-speaking countries and the US. The beliefs uncovered in the content analysis spanned a wide thematic breadth revealing a subjective and heterogeneous consumer understanding of organic food. A total of 65 beliefs regarding 22 superordinate themes has been identified. On a more abstract level, the themes were attributed to four main themes. (1) The first main theme ‘product’ comprises beliefs directly associated with organic food products and its attributes (e.g., related to the themes ‘food safety’, ‘healthiness’, and ‘price’). (2) The second main theme ‘food system’ comprehends beliefs regarding the global strengths and weaknesses of the organic versus the conventional food systems (e.g., related to the themes ‘food security’, ‘production scale’, and ‘farmer welfare’). (3) The third main theme ‘authenticity’ embraces beliefs regarding the meaning and credibility of organic-certified food (e.g., related to the themes ‘organic label’, ‘retail outlet or brand’, and ‘packaging’). (4) The fourth main theme ‘production’ encompasses beliefs regarding the farming and production process and its external impacts (e.g., related to the themes ‘environment’, ‘animal welfare’, and ‘biodiversity’).

A comparison of how frequently the beliefs and topics were mentioned—that is their level of salience—revealed significant differences between the two cultural contexts. ‘Product’ beliefs were frequently mentioned in both regional contexts, but significantly more in the US, where it was the most salient main theme. Also ‘food system’ beliefs were more salient in the US than in German-speaking countries. The most important main theme in the German-speaking countries—and significantly more frequently discussed than in the US—was ‘authenticity’. ‘Production’ was the least-salient main theme in both regions.

These findings deliver an up-to-date and comprehensive picture of consumer perception of organic food. While confirming many beliefs and topics considered in extant literature, *Essay I* also contributes several novel notions unattended by previous literature, such as

beliefs regarding price, biodiversity, animal welfare, working conditions, farmer welfare, organic label, origin, packaging, production scale, and system integrity.

Essay II. As a methodological study, *Essay II* demonstrated the compatibility of qualitative content analysis and neural-network-based semantic filtering using the Universal Sentence Encoder. The Universal Sentence Encoder converted all sentences in the data set as well as all beliefs categories from the content analysis (i.e., input sentences) into sentence embeddings with semantic features. To filter sentences that are similar to the input sentences, *Essay II* suggested several criteria for selecting a similarity threshold. Besides manual inspection of the semantic similarity, *Essay II* proposes comparisons of the total number of sentences filtered as well as correlations between the frequencies per category in the content analysis and the automatic filtering. At the selected similarity threshold, the filtering of text using sentence embeddings from pre-trained neural networks yielded similar results compared to the content analysis. The proposed approach fruitfully combines the scale of computational analysis with the depth of qualitative research. By using a classical content analysis as point of departure, it encounters the criticism of mere NLP analyses lacking theoretical grounding and case sensitivity (Boyd & Crawford, 2012). The method triangulation in *Essay II* finds convergent results from the two text-mining methods, and thus provides validity to using sentence embeddings in the context of organic food, such as in the subsequent *Essay III*, and suggests its application in consumer research in general (Renz et al., 2018; Thurmond, 2001).

Essay III. The third essay shed light on topic or attribute agenda-setting effects between the media and the public agenda using the examples of two major news outlets in Germany and the US for the time period of 2007 to 2020. First, topic modeling uncovered the following ten topics discussed in the media and the public agendas regarding organic food: ‘food safety & chemicals & GMO’, ‘food products & quality’, ‘health & nutrition’, ‘environment & climate change & energy’, ‘farming’, ‘animal welfare & meat consumption’, ‘retailers & prices’, ‘economy & affordability’, ‘politics’, and ‘evidence’. Correlational analyses revealed that the salience of topics in news articles significantly influences the salience of topics in reader comments. Cross-lagged correlations showed that media articles do not only determine the topics discussed in their very own comments, but also comments made to articles of the following two weeks. No such effect in the opposite direction was found, that is, commenters do not influence the content of future

media articles. In addition to these typical agenda-setting metrics, which consider each topic separately, *Essay III* also incorporated agenda diversity and distance metrics from information and probability theory to analyze the dynamics in the distribution of topics on the agenda. These metrics confirmed the close semantic relationship between articles and comments, and, in addition, pointed out specific points in time, where the semantic distance between the media and public agendas was high and where agenda diversity was particularly low. This could be traced back to real world events impacting the agendas, such as general political events and changes in political leadership, as well as (debate on) legal changes and mergers of large retail companies specific to organic food and agriculture. To conclude, *Essay III* made a contribution in being the first to comprehensively document the topic or attribute agenda-setting function of the media in the context of sustainable food consumption.

Essay IV. The fourth essay investigated for the German context whether the relative salience of topics discussed online influenced the priming effects for different organic food-related outcomes: the likelihood of choosing organic over conventional eggs and milk in two choice experiments, and the acceptance of a policy promoting organic agriculture. A preliminary text-mining study had revealed ‘animal welfare’ as a topic of high online salience and ‘biodiversity’ as a topic of low online salience. Subsequently, both topics were employed to prime attitudes toward organic food and affect subsequent behavior in an online experiment. In terms of the methodological implementation, a latent class multinomial logit model was fitted to the choice experiment data. Typically, latent class analysis is destined to identify latent consumer segments (e.g., Jonge et al., 2015). Following previous examples (e.g., Thøgersen & Alfinito, 2020), *Essay IV* showed that priming groups can be imposed as latent classes in order to detect significant differences between groups. The hypotheses could be confirmed with regard to policy acceptance and choices regarding eggs. As hypothesized, the high salience topic was more effective at priming organic food behavior than the low salience topic. Priming effectiveness further depended on prime strength, that is situational salience, which was varied by either only hinting or strongly pointing toward the two topics. The priming effects were positively moderated by attitude strength and/or experience with buying organic food. However, results regarding milk choices were counter to hypothesized. The likelihood of choosing organic milk was higher in the control group than in the priming groups. The different findings for the two products were unexpected and explanations are only speculative.

Participants had been screened to be egg and milk buyers. Both eggs and milk belong to the most purchased organic products in Germany and are also among the first ones when consumers initially start buying organic food (Juhl et al., 2017). Although both animal welfare and biodiversity are important reasons to buy organic food in Germany (Federal Office for Agriculture and Food, 2020), participants could possibly have perceived neither of the topics as relevant for milk purchases and thus not have assimilated their behavior to the prime (DeCoster & Claypool, 2004). Personal associations as opposed to the priming topics could have been more salient in the moment of milk choice, as milk is a habitual choice and milk products are purchased slightly more frequently than eggs. Further research is needed on priming based on online salience, in particular with regard to product-specific effects. However, *Essay IV* produced novel insights regarding the link between the consumer voice in the online sphere and judgment and decision-making of the general adult population in Germany in the context of sustainable food consumption.

UGC as data for consumer research. This thesis exemplifies the potential of UGC for consumer research. Compared to data from social networks or product reviews, online comments from forums and news websites—as used in this dissertation—have yet been little exploited for consumer insight. The presented studies consistently revealed that UGC is a meaningful and valid data source to learn about what consumers think and do. The analysis of UGC revealed detailed knowledge about beliefs and topics discussed online and their respective degree of salience. Text mining proved suitable for research into credence goods like organic food and to uncover subjective meanings considering consumers' own terminology. This thesis explored different types of online text data (reader comments, forum comments, news articles) as well as data in German and English language. It applied different, both established and novel, text-mining techniques (content analysis, topic modeling based on LDA and topic modeling based on clustered sentence embeddings), different units of analysis (words, sentences, paragraphs), and different units of measurement (frequency of belief or topic occurrence, topic distribution, and share of comments referencing a topic). The different text-mining methods serve different consumer research purposes. While the neural-network and the LDA topic modeling (*Essays III* and *IV*) delivered topics at a higher abstraction level (8 and 18 topics), the unique advantage of the content analysis (*Essay I*) is its level of detail and the integration of extant consumer research evidence (65 beliefs regarding 22 topics). While still counting the frequency of mention (i.e., salience), the content analysis took stock of all

beliefs. Thus, in contrast to topic modeling techniques, which mark out the most prevalent topics, manual content analysis could uncover subtle notions and differences in beliefs, which are—irrespective of their degree of salience—important for consumer research to understand consumers' rationales.

Moreover, the dissertation investigated relationships between text data, showing the topic relationship between news articles and reader comments, as well as the reflection of online salience in attitudinal and choice data from the general population.

As noted in Section 1.3 and demonstrated throughout the thesis, UGC can be a main or complementary data source, which especially lends itself for both snapshot-type and in-depth analysis of current and long-term consumer perception. Moreover, analyses can be continuously updated and thus allow for time analyses at little cost. It can be noted that a large part of sustainable and organic food consumption research relies on data from surveys, interviews, and focus groups, which can be biased in several ways (see Section 1.3). Furthermore, survey data are very structured and do not allow for detecting new aspects of consumer perception. In contrast, qualitative interview and focus group data are indeed exploratory, but limited to small samples. UGC can provide an effective means to tap the voices of a large number of consumers and to uncover latent needs and understandings. In conclusion, UGC could thus further advance research into sustainable consumption and eventually contribute to a faster transition toward a green economy.

Reflections on triangulation, validity and behavioral relevance. This thesis performed data, investigator, and method triangulation within and across the four essays to increase the validity and the interpretative quality of the results (Denzin, 1978). An overview will be provided in the following (see also Table 1).

First, comparing the results on salient beliefs and topics across the four consecutive essays, the internal validity of the text-mining results can be confirmed: Although on different levels of abstraction (i.e., number of topics), conceptually consistent findings regarding the identified beliefs and topics across different methods as well as small and large data sets safeguard the interpretability and trustworthiness of the findings. Thus, convergent validity is given as multiple measures converge to the same concept. This method and data triangulation indicates that the selected measures accurately reflect the constructs measured (Denzin, 1978; Renz et al., 2018). The different methods were content analysis, filtering based on semantic similarity, topic modeling based on LDA,

and topic modeling based on clustered sentence embeddings. The latter and rather novel method also proved valid. The topics (i.e., clusters of sentence embeddings) were more semantically coherent (*Essay III*) compared to the topics identified by the topic modeling based on the established LDA algorithm (*Essay IV*), but the topics itself were conceptually comparable. In addition, the identified beliefs and topics reflected what was reported by previous research on UGC on organic food (Lyu & Choi, 2020; Meza & Park, 2016; Olson, 2017) and thus suggest concurrent validity. However, previous literature on beliefs and topics of organic food in online discussion is still scarce and the depth and scope of the presented analyses on UGC on organic food is unprecedented. Moreover, for *Essay I*, investigator triangulation, that is, the use of several coders to categorize the text, assured the reliability of the qualitative content analysis. This was confirmed by the calculation of Cohen's Kappa for intercoder reliability.

Second, the text-mining results have content validity and are generalizable. The beliefs and topics identified are consistent with the findings of previous qualitative and survey research (Hemmerling et al., 2015; Hughner et al., 2007) and thus provide an adequate coverage of the target content. In addition, the detailed analysis of *Essay I* extended the literature's view on organic food perception by previously neglected beliefs and topics salient to consumers. Nevertheless, the sample of the online population and online populations in general might still partly differ from the general population (Berger et al., 2020), although this dissertation chose main-stream and easy-access sources to ensure a diverse audience. It is therefore advisable to perform further checks on the representativeness of insights from UGC for the general population.

Third, *Essay III* provides convergent validity for the identified the topic agenda-setting between the media and the public as several agenda-setting metrics (topic correlation, agenda distance, and agenda diversity) testify a close relationship between the topics discussed in articles and comments.

Fourth, *Essay IV* tested and largely confirmed that online salience corresponds to priming effectiveness on sustainable consumer behavior. Despite the need for further research, the external validity, that is the behavioral relevance, of the salience of beliefs and topics voiced online has become evident. The dissertation took a novel perspective on the relationship between the online and offline sphere. Respective research in the context of sustainable consumption is scarce. An exception is the study of Lyu and Choi (2020).

Despite not analyzing consumer behavior, they show a connection between the online and offline sphere by finding that positively evaluated attributes mentioned online are closely related to survey results on product attributes responsible for product satisfaction. In contrast, the external relevance of online comments on news websites has been little studied. Previous research mostly focused on the relationship between product reviews and sales. Furthermore, these studies mostly analyze review volume and valence, but not the actual content (i.e., topics) and its implications for the products chosen (for a meta-analytic study of 96 studies on electronic word-of-mouth, see, Babić Rosario et al., 2016). Instead this research did not only analyze the frequency of comments and its influence on revealed or stated behavior, but also accounted for the effect of the actual content of comments (i.e., discussion topics) and the relative weight of different topics in the discussion. Thus, this dissertation contributes to evidence for a connection between the online and offline sphere in the context of sustainable consumption by showing that topics salient online can be employed to promote related consumer behavior offline.

4.2 Research implications

This thesis contributes to organic food, sustainable consumption, and research. It entails several implications for future research, in particular, regarding the elicitation of beliefs in survey research, the use of UGC and different text-mining approaches for its analysis, and research into the interplay between online salience of topics and its relevance for consumer behavior.

Implications for belief elicitation in survey research. The text-mining studies revealed a large diversity of beliefs and topics, going beyond beliefs and topics reported in extant literature and the breadth and number of belief items commonly elicited in surveys. In particular, *Essay I* entails several implications for more reliable and valid measurement of consumer beliefs that reflects consumer perception of organic food more comprehensively: Future survey research should elicit a larger number of items, positively, neutrally, and negatively phrased items, as well as belief items regarding both organic and conventional food. Furthermore, beliefs held simultaneously have to be considered. The belief phrasings in the presented category system closely reflect the terminology employed by the commenters. Using consumers' phrasing to build survey

items has been shown to benefit the reliability and validity of measurements (Eisner et al., 2019). Finally, the country comparisons indicated that future research should consider the differences in consumer perception of organic food when selecting items to elicit consumer beliefs about organic food. Previous studies have often elicited not only few items to characterize consumer perception but also very similar items across cultural contexts. Given these insights from text mining, future research in the German and US contexts can adopt a more country-specific set of items and thus represent consumer beliefs more accurately. Moreover, the in-depth results yielded suggest the application of text mining of UGC to other areas of sustainable consumption to yield a differentiated picture of consumer perception.

Implications for text mining. This thesis applied different text mining methods with converging results regarding the identified topics and their salience. While qualitative content analysis is only recommended for smaller samples due to required time and human resources, large data samples can be analyzed with the widely applied LDA algorithm for topic modeling. Another large-scale topic modeling technique is the clustering of neural network-based sentence embeddings according to semantic similarity (*Essay III*). Although also using off-the-shelf neural network solutions like USE or XLING, the latter approach is computationally more intense than LDA topic modeling. However, it safeguards for the consideration of the semantic context of words and sentences and thus generates semantically more coherent topics. In addition, sentence embedding-based approaches can be applied in a multilingual fashion as shown in *Essay III*. Moreover, as a methodological study, *Essay II* provides implications for effectively combining and validating small and large-scale text-analysis methods. The comparison of results from the content analysis and the embedding-based filtering yielded similar results, and thus provided validity for a combined approach suggested for subsequent implementation: The embedding-based filtering could next be applied to an unseen, arbitrarily large data set on organic food to gain insight into the distribution and salience of beliefs of a larger sample. Furthermore, the suggested approach can be applied multilingually by taking input sentences from one language and filtering sentences in a different language. This feature could serve research into cross-cultural comparisons. This dissertation made a first case for the of combination of two small and large-scale text-mining techniques and demonstrated different text analyses techniques across the

essays. Future research should further investigate and compare the coherence, case suitability, and compatibility of different text-mining approaches.

Implications for agenda-setting research. The thesis provided empirical evidence for the media setting the topics on the public's agenda in the context of organic food. Future research could extend the insights for further examples of sustainable consumption. In a subsequent step to the topic agenda-setting shown, future research could investigate media framing or sentiment agenda-setting effects (Vargo et al., 2014), that is whether and how the media also influence the sentiment regarding organic food or sustainable consumption in general (i.e., whether a topic is evaluated positively, neutrally, or negatively). This would show that media are not only a source of knowledge but also meaningful in the formation of consumer attitude toward organic food. Moreover, the agenda-setting insights from this dissertation are limited to two major news outlets and their reader audience. Further research could include more media outlets and, in addition, explore intermedia agenda-setting, that is influence between platforms (Lim, 2011). Given the increasing fragmentation of the media market (Takeshita, 2006), the influence between several news media and social media platforms could be investigated.

Implications for research into the interplay of online salience and consumer behavior. The thesis also aimed at exploring the effectiveness of priming based on online salience and its potential application in the communication of sustainable consumption. As expected the topic with high online salience was more effective compared to the low online salience topic in priming eggs choices and policy acceptance. Results on milk choices were inconclusive. Further research efforts should consider more products and product categories to further investigate how online salience translates to consumer behavior and sustainable consumption behavior in particular. Nevertheless, the results of this thesis indicated the potential of UGC to activate product-related knowledge and attitudes and can be used to promote desired behaviors. While the two selected priming topics (animal welfare and biodiversity) are rated of similar importance by consumers (Federal Office for Agriculture and Food, 2020), the detected differences in online salience of these topics was considerably larger and made a difference when implemented as prime in the experiment. This showed that considering salience and not only importance matters for explaining consumer behavior. Finally, it remains to be noted that self-reported behaviors and choice experiments, such as applied in *Essay IV*, witness

hypothetical behavior. Previous research has discussed how such measures often do not match with actual purchase behavior given the attitude-behavior gap (Carrington et al., 2010; Janssen, 2018). This critique applies in particular to purchase decisions with an ethical dimension, where participants are prone to give socially desirable answers, such as sustainable behavior (Auger & Devinney, 2007; Prothero et al., 2011) and organic food consumption in particular (Padel & Foster, 2005). Digital transformation of communication and consumption is a dynamic and ongoing process. Further and continuous exploration is needed to understand the links between the online sphere and actual sustainable consumption behavior (Frick et al., 2020).

4.3 Managerial and policy implications

The findings from this dissertation provide practical implications for different players such as marketing and communication practitioners, policy makers, and media outlets. The implications concern insights on consumer perception, respective regional differences as well as subjectivity, and the interplay between online salience and consumption behavior. Furthermore, practical implications can be drawn for the role of the media. The findings and implications from the case of organic food can inform other areas of sustainable consumption.

Consumer perception. The text mining has shown that UGC delivers insights into current consumer perception and behavior regarding sustainable consumption using the example of organic food. The different analyses of UGC conveyed a detailed picture of what is on consumers' agendas and the relative salience of beliefs and topics related to organic food. Text mining could be used for continuous market monitoring. Marketing communication can employ these insights to match products and their presentation with the topics salient to consumers and their expectations. Knowing consumers' expectations and the degree to which they believe in the promoted benefits is key for successful marketing as well as the introduction of new products into the market (Lusk et al., 2014). Although more research on different products and product categories is needed, the thesis demonstrated the employability of online salience for effective marketing or policy communication. Results suggested that the salience of topics is an important aspect of the 'choice architecture' for consumer choices, and that salience-based priming (Reisch

& Zhao, 2017) can be an effective “nudge” (Thaler & Sunstein, 2009). This means that marketers can use online salient topics to guide consumers toward more sustainable behaviors. A review of salience and priming interventions to nudge people toward healthier choices (i.e., providing relevant health information) revealed mixed effectiveness of mere salience nudges (Wilson et al., 2016). The authors propose the combination with other nudges or interventions. Salience nudges seem to become more effective when combined with, for example, modifying the visibility, availability, and/or accessibility of choices.

Finally, this thesis suggests that the degree of online salience of topics regarding organic food cannot only influence judgment and decision-making with regard to market products but also with regard to political behavior. Topics with high salience are more likely to make people vote for policies supporting organic agriculture. Thus, for public opinion polling, monitoring the online discussion can be an important indicator of political opinion.

Country differences in consumer perception. This dissertation shed light on consumer perception of organic food in Germany and the US. Despite commonalities and although food markets are more and more connected, it was found that consumer perception differs across these two cultural contexts. This research revealed which beliefs and topics are more prevalent in Germany and the US. In brief, US commenters focused on food safety and quality issues, while German-speaking commenters made emphasis on the authenticity of products and differences in standards and labels. The organization of the organic food markets in the two regions reflects the different foci of attention in consumer perception. Both investigated regions count on a solid legal framework for the production of organic food. In the US, mainly the USDA organic standard and the corresponding label are proliferated. Moreover, choosing organic options is considered a way to avoid GMO-food. In contrast, in Germany, there are several private organic labels, which represent stricter standards compared to the national and EU organic standards. The differentiated standard and labeling landscape could partly explain the belief in different degrees of “organicalness”. The insights on consumer perception delivered by this thesis provide leverage for the further proliferation of organic farming and the promotion of organic food sales. Marketers entering or operating in these countries are advised to consider in their communication these perceptual differences, which have been shown

for the case of organic food, but are likely to apply to other areas of sustainable consumption.

Subjective versus objective organic food benefits. The text-mining analyses have shown that consumer perception is very subjective and does not always match with the actual benefits of organic food production and consumption. For example, a substantial share of especially German commenters believed in different degrees of “organicalness”, i.e., products were perceived as more or less organic depending on the organic label, the country or region of origin, the retail outlet or brand, the type of packaging, and even the product category. On this matter, policy makers involved in the development of food standards as well as marketers are suggested to deliver more transparent information on actual standards and benefits and thus contribute to more evidence-based comparison of products and consumer decisions. Moreover, marketers can rectify consumers’ expectations in case of exaggerated expectations and thus safeguard against public deception (Thorsøe, 2015). For example,—besides enhanced animal welfare in organic animal husbandry—assumed perceived health benefits of organic over conventional food remain the key argument for organic food purchases in Germany (Federal Office for Agriculture and Food, 2020). In terms of health benefits, organic food consumption reduces the exposure to pesticide residues and antibiotic-resistant bacteria, however, organic food is unlikely to be more nutritious than conventional food (Mie et al., 2017; Smith-Spangler et al., 2012). Communication should focus on core benefits of organic farming and products, as a mismatch between actual benefits and consumer expectations threatens the trust in the organic food system (Thorsøe, 2015).

Further, it can be stated that the online salience of topics differs and does not always mirror the pressing problems. For example, an update of the planetary boundary framework has identified the integrity of the biosphere and climate change as two core planetary boundaries out of a total nine, the transgression of which alone could push the earth system persistently into a new state threatening its resilience (Steffen et al., 2015). However, biodiversity losses due to farming practices and respective advantages of organic farming were little discussed online. As stated above and documented in literature, consumer perception of organic food is subjective and not completely evidence-based. It is difficult for consumers to judge the actual impact of different ecologically-relevant behaviors, such as consuming organic food, reducing meat

consumption or avoiding air-transported foods (Jungbluth et al., 2000). Consumer perception does not always match with the actual impact. For example, Tobler et al. (2011) have found that consumers overrate the environmental benefits from avoiding excessive packaging, and underestimate the environmental benefits of eating organically-produced food or reducing meat consumption. It is therefore proposed that control bodies and communication practitioners further invest in information provision and the sharpening consumer perception of organic products and farming, as well as sustainable consumption practices in general.

The role of the media. Germans spent a daily average of 22 minutes reading news media (SevenOne Media, 2020), with digital newspapers attaining almost the same reach as print (ZMG, 2020). The media are an important source of knowledge and media consumption does not remain without effects. This dissertation displayed that the media affect consumer perception of organic food. Considering the provided insights on agenda setting, the news media can be highlighted as a decisive player and tool to influence consumers and shape their understanding of global societal issues and possible avenues to more sustainable consumption. Thus, this also assigns responsibility to media outlets as agenda setters. Moreover, the media not only impact their reader audience—as shown in this dissertation—, but has also been shown to determine corporate environmental responsibility agendas (Pollach, 2014). Also consumers' online voices do not remain unheard—although this thesis provided no evidence for reader comments affecting the media agenda—consumers' social media comments and reactions to corporate actions have been shown to affect corporate communication in return, as Ruggeri and Samoggia (2018) have demonstrated for companies along the palm-oil value chain.

Furthermore, the media's power to impact the agendas of their readers offers potential opportunities to policy makers, which can use the media as a channel to influence consumer perception of sustainable consumption. Policy measures, such as the introduction, adaptation, and promotion of sustainability standards, can be accompanied by media campaigns, and thus raise and increase the awareness, familiarity, and salience with consumers. Furthermore, research has attested that consumers particularly pay attention to negative news reports (Müller & Gaus, 2015; Yadavalli & Jones, 2014). Müller and Gaus (2015) showed that negative media reports influenced intention, attitude

and trust regarding organic food. Thus, media campaigns could be crucial to maintain or reestablish trust in sustainability standards and labels.

As the detrimental consequences of current consumption patterns become more evident and digitization offers additional visibility of consumer reasoning, research into consumer perception and behavior in the online sphere and its influencing factors and implications is key to building a more sustainable economy and society.

References

- Aarset, B., Beckmann, S., Bigne, E., Beveridge, M., Bjorndal, T., Bunting, J., . . . Young, J. (2004). The European consumers' understanding and perceptions of the "organic" food regime. *British Food Journal*, *106*(2), 93–105. <https://doi.org/10.1108/00070700410516784>
- Abdi-Herrle, S. (2018). *Mediale Themensetzung in Zeiten von Web 2.0. Politische Kommunikation und demokratische Öffentlichkeit: Band 16*. Baden-Baden: Nomos Verlagsgesellschaft.
- Abrams, K. M., Meyers, C. A., & Irani, T. A. (2010). Naturally confused: Consumers' perceptions of all-natural and organic pork products. *Agriculture and Human Values*, *27*(3), 365–374. <https://doi.org/10.1007/s10460-009-9234-5>
- Adamowicz, W., Bunch, D., Cameron, T. A., Dellaert, B. G. C., Hanneman, M., Keane, M., . . . Swait, J. (2008). Behavioral frontiers in choice modeling. *Marketing Letters*, *19*(3-4), 215–228. <https://doi.org/10.1007/s11002-008-9038-1>
- Ader, C. R. (1995). A Longitudinal Study of Agenda Setting for the Issue of Environmental Pollution. *Journalism & Mass Communication Quarterly*, *72*(2), 300–311. <https://doi.org/10.1177/107769909507200204>
- Aertsens, J., Verbeke, W., Mondelaers, K., & van Huylenbroeck, G. (2009). Personal determinants of organic food consumption: A review. *British Food Journal*, *111*(10), 1140–1167. <https://doi.org/10.1108/00070700910992961>
- Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, *50*(2), 179–211. [https://doi.org/10.1016/0749-5978\(91\)90020-T](https://doi.org/10.1016/0749-5978(91)90020-T)
- Ajzen, I. (2002). Constructing a Theory of Planned Behavior Questionnaire. Retrieved from <http://www.people.umass.edu/ajzen/tpb.html>*, (accessed November 12, 2020).
- Ajzen, I., & Fishbein, M. (1980). *Understanding attitudes and predicting social behavior* (Pbk. ed.). Englewood Cliffs, N.J: Prentice-Hall.
- Ajzen, I., & Fishbein, M. (2005). The influence of attitudes on behavior. In B. T. Johnson, D. Albarracin, & M. P. Zanna (Eds.), *The Handbook of Attitudes* (pp. 173–221). Mahwah, N.J: Lawrence Erlbaum Associates.
- Alphonse, R., Alfnes, F., & Sharma, A. (2014). Consumer vs. citizen willingness to pay for restaurant food safety. *Food Policy*, *49*, 160–166. <https://doi.org/10.1016/j.foodpol.2014.06.009>
- Aschemann-Witzel, J., & Niebuhr Aagaard, E. M. (2014). Elaborating on the attitude-behaviour gap regarding organic products: Young Danish consumers and in-store food choice. *International Journal of Consumer Studies*, *38*(5), 550–558. <https://doi.org/10.1111/ijcs.12115>

- Auger, P., & Devinney, T. M. (2007). Do what consumers say matter? The misalignment of preferences with unconstrained ethical intentions. *Journal of Business Ethics, 76*(4), 361–383. <https://doi.org/10.1007/s10551-006-9287-y>
- Babić Rosario, A., Sotgiu, F., Valck, K. de, & Bijmolt, T. H. (2016). The Effect of Electronic Word of Mouth on Sales: A Meta-Analytic Review of Platform, Product, and Metric Factors. *Journal of Marketing Research, 53*(3), 297–318. <https://doi.org/10.1509/jmr.14.0380>
- Balducci, B., & Marinova, D. (2018). Unstructured data in marketing. *Journal of the Academy of Marketing Science, 46*(4), 557–590. <https://doi.org/10.1007/s11747-018-0581-x>
- Baroni, L., Cenci, L., Tettamanti, M., & Berati, M. (2007). Evaluating the environmental impact of various dietary patterns combined with different food production systems. *European Journal of Clinical Nutrition, 61*(2), 279–286. <https://doi.org/10.1038/sj.ejcn.1602522>
- Bazzani, C., Caputo, V., Nayga, R. M., & Canavari, M. (2017). Revisiting consumers' valuation for local versus organic food using a non-hypothetical choice experiment: Does personality matter? *Food Quality and Preference, 62*, 144–154. <https://doi.org/10.1016/j.foodqual.2017.06.019>
- Bazzani, C., Gustavsen, G. W., Nayga, R. M., & Rickertsen, K. (2018). A comparative study of food values between the United States and Norway. *European Review of Agricultural Economics, 45*(2), 239–272. <https://doi.org/10.1093/erae/jbx033>
- Bellotti, E., & Panzone, L. (2016). Media effects on sustainable food consumption. How newspaper coverage relates to supermarket expenditures. *International Journal of Consumer Studies, 40*(2), 186–200. <https://doi.org/10.1111/ijcs.12242>
- Berger, I. E., & Mitchell, A. A. (1989). The Effect of Advertising on Attitude Accessibility, Attitude Confidence, and the Attitude-Behavior Relationship. *Journal of Consumer Research, 16*(3), 269–279. <https://doi.org/10.1086/209213>
- Berger, J., Humphreys, A., Ludwig, S., Moe, W. W., Netzer, O., & Schweidel, D. A. (2020). Uniting the Tribes: Using Text for Marketing Insight. *Journal of Marketing, 84*(1), 1–25. <https://doi.org/10.1177/0022242919873106>
- Bettman, J. R. (1979). Memory Factors in Consumer Choice: A Review. *Journal of Marketing, 43*(2), 37–53. <https://doi.org/10.1177/002224297904300205>
- Bezawada, R., & Pauwels, K. (2013). What Is Special About Marketing Organic Products? How Organic Assortment, Price, and Promotions Drive Retailer Performance. *Journal of Marketing, 77*(1), 31–51. <https://doi.org/10.1509/jm.10.0229>
- Birau, M. M., & Faure, C. (2018). It is easy to do the right thing: Avoiding the backfiring effects of advertisements that blame consumers for waste. *Journal of Business Research, 87*, 102–117. <https://doi.org/10.1016/j.jbusres.2018.02.026>
- Bitsch, V., Koković, N., & Rombach, M. (2014). Risk Communication and Market Effects during Foodborne Illnesses: A Comparative Case Study of Bacterial Outbreaks in the U.S. and in Germany. *International Food and Agribusiness Management Review, 17*(3), 97–114. <https://doi.org/10.22004/AG.ECON.183451>

- Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM*, 55(4), 77–84. <https://doi.org/10.1145/2133806.2133826>
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research*, 3(Jan), 993–1022.
- Blumenthal-Barby, J. S., & Burroughs, H. (2012). Seeking better health care outcomes: the ethics of using the "nudge". *The American Journal of Bioethics : AJOB*, 12(2), 1–10. <https://doi.org/10.1080/15265161.2011.634481>
- Börjesson Rivera, M., Håkansson, C., Svenfelt, Å., & Finnveden, G. (2014). Including second order effects in environmental assessments of ICT. *Environmental Modelling & Software*, 56, 105–115. <https://doi.org/10.1016/j.envsoft.2014.02.005>
- Bowen, G. A. (2009). Document analysis as a qualitative research method. *Qualitative Research Journal*, 9(2), 27–40. <https://doi.org/10.3316/QRJ0902027>
- Boyd, D., & Crawford, K. (2012). Critical Questions for Big Data. *Information, Communication & Society*, 15(5), 662–679. <https://doi.org/10.1080/1369118X.2012.678878>
- Boydston, A. E., Bevan, S., & Thomas, H. F. (2014). The Importance of Attention Diversity and How to Measure It. *Policy Studies Journal*, 42(2), 173–196. <https://doi.org/10.1111/psj.12055>
- Branthwaite, A., & Patterson, S. (2011). The power of qualitative research in the era of social media. *Qualitative Market Research: an International Journal*, 14(4), 430–440. <https://doi.org/10.1108/13522751111163245>
- Bray, J., Johns, N., & Kilburn, D. (2011). An exploratory study into the factors impeding ethical consumption. *Journal of Business Ethics*, 98(4), 597–608. <https://doi.org/10.1007/s10551-010-0640-9>
- Brehm, J. W. (1966). *A theory of psychological reactance*. New York: Academic Press.
- Brún, A. de, McCarthy, M., McKenzie, K., & McGloin, A. (2014). Weight stigma and narrative resistance evident in online discussions of obesity. *Appetite*, 72, 73–81. <https://doi.org/10.1016/j.appet.2013.09.022>
- Büschken, J., & Allenby, G. M. (2016). Sentence-Based Text Analysis for Customer Reviews. *Marketing Science*, 35(6), 953–975. <https://doi.org/10.1287/mksc.2016.0993>
- Camacho-Otero, J., Boks, C., & Pettersen, I. N. (2019). User acceptance and adoption of circular offerings in the fashion sector: Insights from user-generated online reviews. *Journal of Cleaner Production*, 231, 928–939. <https://doi.org/10.1016/j.jclepro.2019.05.162>
- Carrington, M. J., Neville, B. A., & Whitwell, G. J. (2010). Why ethical consumers don't walk their talk: Towards a framework for understanding the gap between the ethical purchase intentions and actual buying behaviour of ethically minded consumers. *Journal of Business Ethics*, 97(1), 139–158. <https://doi.org/10.1007/s10551-010-0501-6>
- Cascio, C. N., O'Donnell, M. B., Bayer, J., Tinney, F. J., & Falk, E. B. (2015). Neural Correlates of Susceptibility to Group Opinions in Online Word-of-Mouth

- Recommendations. *Journal of Marketing Research*, 52(4), 559–575.
<https://doi.org/10.1509/jmr.13.0611>
- Cer, D., Yang, Y., Kong, S., Hua, N., Limtiaco, N., St. John, R., . . . Kurzweil, R. (2018). Universal Sentence Encoder. *ArXiv*. Retrieved from <http://arxiv.org/pdf/1803.11175v2>
- Cerri, J., Thøgersen, J., & Testa, F. (2019). Social desirability and sustainable food research: A systematic literature review. *Food Quality and Preference*, 71, 136–140. <https://doi.org/10.1016/j.foodqual.2018.06.013>
- Chatzidakis, A., & Mitussis, D. (2007). Computer ethics and consumer ethics: the impact of the internet on consumers' ethical decision-making process. *Journal of Consumer Behaviour*, 6(5), 305–320. <https://doi.org/10.1002/cb.223>
- Chekima, B., Oswald, A. I. @., Wafa, S. A. W. S. K., & Chekima, K. (2017). Narrowing the gap: Factors driving organic food consumption. *Journal of Cleaner Production*, 166, 1438–1447. <https://doi.org/10.1016/j.jclepro.2017.08.086>
- Chen, M.-F. (2020). Selecting Environmental Psychology Theories to Predict People's Consumption Intention of Locally Produced Organic Foods. *International Journal of Consumer Studies*, 44(5), 455–468. <https://doi.org/10.1111/ijcs.12578>
- Chidambaram, M., Yang, Y., Cer, D., Yuan, S., Sung, Y.-H., Strophe, B., & Kurzweil, R. (2019). Learning Cross-Lingual Sentence Representations via a Multi-task Dual-Encoder Model. In *Proceedings of the 4th Workshop on Representation Learning for NLP (ReplANLP-2019)*. Retrieved from <https://arxiv.org/pdf/1810.12836>
- Cho, S., & Hong, Y. (2009). Netizens' evaluations of corporate social responsibility: Content analysis of CSR news stories and online readers' comments. *Public Relations Review*, 35(2), 147–149. <https://doi.org/10.1016/j.pubrev.2008.09.012>
- Christensen, K., Liland, K. H., Kvaal, K., Risvik, E., Biancolillo, A., Scholderer, J., . . . Næs, T. (2017). Mining online community data: The nature of ideas in online communities. *Food Quality and Preference*, 62, 246–256. <https://doi.org/10.1016/j.foodqual.2017.06.001>
- Chwialkowska, A., & Flicinska-Turkiewicz, J. (2020). Overcoming perceived sacrifice as a barrier to the adoption of green non-purchase behaviours. *International Journal of Consumer Studies*, 45(2), 205–220. <https://doi.org/10.1111/ijcs.12615>
- Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, 82(6), 407–428. <https://doi.org/10.1037/0033-295X.82.6.407>
- Conway-Silva, B. A., Filer, C. R., Kenski, K., & Tsetsi, E. (2018). Reassessing Twitter's Agenda-Building Power. *Social Science Computer Review*, 36(4), 469–483. <https://doi.org/10.1177/0894439317715430>
- Cooper, G., Green, N., Burningham, K., Evans, D., & Jackson, T. (2012). Unravelling the Threads: Discourses of Sustainability and Consumption in an Online Forum. *Environmental Communication*, 6(1), 101–118. <https://doi.org/10.1080/17524032.2011.642080>

- Costanigro, M., Deselnicu, O., & Kroll, S. (2015). Food beliefs: Elicitation, estimation and implications for labeling policy. *Journal of Agricultural Economics*, 66(1), 108–128. <https://doi.org/10.1111/1477-9552.12085>
- Danner, H., & Menapace, L. (2020). Using Online Comments to Explore Consumer Beliefs Regarding Organic Food in German-Speaking Countries and the United States. *Food Quality and Preference*, 83(103912). <https://doi.org/10.1016/j.foodqual.2020.103912>
- Darby, M. R., & Karni, E. (1973). Free competition and the optimal amount of fraud. *The Journal of Law and Economics*, 16(1), 67–88. <https://doi.org/10.1086/466756>
- DataReportal. (2020). Digital 2020: July Global Statshot Report. Retrieved from https://www-statista-com.eaccess.ub.tum.de/statistics/269615/social-network-penetration-by-region/*, (accessed December 3, 2020).
- DeCoster, J., & Claypool, H. M. (2004). A meta-analysis of priming effects on impression formation supporting a general model of informational biases. *Personality and Social Psychology Review*, 8(1), 2–27. https://doi.org/10.1207/S15327957PSPR0801_1
- Denham, B. E. (2010). Toward Conceptual Consistency in Studies of Agenda-Building Processes: A Scholarly Review. *Review of Communication*, 10(4), 306–323. <https://doi.org/10.1080/15358593.2010.502593>
- Denzin, N. K. (1978). *Sociological methods: A sourcebook* (2nd ed.). New York: McGraw-Hill. Retrieved from <http://worldcatlibraries.org/wcpa/oclc/3516374>
- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *ArXiv*. Retrieved from <https://arxiv.org/pdf/1810.04805>
- Dillman, D. A., Smyth, J. D., & Christian, L. M. (2014). *Internet, Phone, Mail, and Mixed-Mode Surveys: The Tailored Design Method* (4th ed.). s.l.: Wiley. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=827492>
- Ditlevsen, K., Denver, S., Christensen, T., & Lassen, J. (2020). A taste for locally produced food - Values, opinions and sociodemographic differences among 'organic' and 'conventional' consumers. *Appetite*, 147, 104544. <https://doi.org/10.1016/j.appet.2019.104544>
- Eisner, L., Clémence, A., Roberts, C., Joost, S., & Theler, J.-M. (2019). Developing Attitude Measures Based on Respondents' Representations of Unfamiliar Objects: An Application to Attitudes toward Biodiversity. *Field Methods*, 31(1), 56–75. <https://doi.org/10.1177/1525822X18797280>
- Ellison, B., Bernard, J. C., Paukett, M., & Toensmeyer, U. C. (2016). The influence of retail outlet and FSMA information on consumer perceptions of and willingness to pay for organic grape tomatoes. *Journal of Economic Psychology*, 55, 109–119. <https://doi.org/10.1016/j.joep.2016.05.002>

- European Commission. (2020a). Organics at a glance. Retrieved from https://ec.europa.eu/info/food-farming-fisheries/farming/organic-farming/organics-glance_en*, (accessed October 19, 2020).
- European Commission. (2020b). Questions and Answers - European Green Deal: Commission prepares new initiatives to boost the organic farming sector. Retrieved from https://ec.europa.eu/commission/presscorner/detail/en/QANDA_20_1539*, (accessed November 5, 2020).
- EUROSTAT. (2018). Households - Level of Internet Access. Retrieved from http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=isoc_ci_in_h&lang=en*, (accessed January 14, 2019).
- EUROSTAT. (2020a). Organic Farming Statistics. Retrieved from https://ec.europa.eu/eurostat/statistics-explained/index.php/Organic_farming_statistics*, (accessed November 5, 2020).
- EUROSTAT. (2020b). Share of households with internet access in Germany from 2007 to 2019. Retrieved from https://www-statista-com//statistics/377677/household-internet-access-in-germany/*, (accessed September 28, 2020).
- Evangelidis, I., & van Osselaer, S. M. (2018). Points of (dis)parity: Expectation disconfirmation from common attributes in consumer choice. *Journal of Marketing Research*, 55(1), 1–13. <https://doi.org/10.1509/jmr.15.0233>
- Fazio, R. H. (1995). Attitudes as object-evaluation associations: Determinants, consequences. In R. E. Petty & J. A. Krosnick (Eds.), *Attitude strength: Antecedents and consequences* (pp. 247–282). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology*, 50(2), 229–238. <https://doi.org/10.1037/0022-3514.50.2.229>
- Fazio, R. H. (1986). How Do Attitudes Guide Behavior? In R. M. Sorrentino & E. T. Higgins (Eds.), *Handbook of motivation and cognition: Foundations of social behavior* (pp. 204–243). New York: Guilford Press.
- Fazio, R. H. (1990). Multiple Processes by which Attitudes Guide Behavior: The Mode Model as an Integrative Framework. In A. Unknown (Ed.), *Advances in Experimental Social Psychology: Vol. 23. Advances in Experimental Social Psychology: Volume 23* (1st ed., Vol. 23, pp. 75–109). s.l.: Elsevier textbooks.
- Fazio, R. H. (2007). Attitudes as Object-Evaluation Associations of Varying Strength. *Social Cognition*, 25(5), 603–637. <https://doi.org/10.1521/soco.2007.25.5.603>
- Fazio, R. H., & Olson, M. A. (2003). Implicit measures in social cognition. research: their meaning and use. *Annual Review of Psychology*, 54, 297–327. <https://doi.org/10.1146/annurev.psych.54.101601.145225>
- Fazio, R. H., & Zanna, M. P. (1981). *Direct Experience And Attitude-Behavior Consistency* (Vol. 14): Elsevier.
- Federal Office for Agriculture and Food. (2020). *Ökobarometer 2019*. Bonn.
- Feldmann, C., & Hamm, U. (2015). Consumers' perceptions and preferences for local food: A review. *Food Quality and Preference*, 40, 152–164. <https://doi.org/10.1016/j.foodqual.2014.09.014>

- Fernqvist, F., & Ekelund, L. (2014). Credence and the effect on consumer liking of food – A review. *Food Quality and Preference*, *32*, 340–353.
<https://doi.org/10.1016/j.foodqual.2013.10.005>
- FiBL. (2020). Data on organic agriculture worldwide 2018. The Statistics.FiBL.org website maintained by the Research Institute of Organic Agriculture (FiBL). Retrieved from http://statistics.fibl.org/world.html*, (accessed October 15, 2020).
- Fishbein, M. (1963). An investigation of the relationship between beliefs about an object and the attitude toward that object. *Human Relations*, *16*, 233–239.
- Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention and behavior: An introduction to theory and research*. Reading: Addison-Wesley.
- Frick, V., & Matthies, E. (2020). Everything is just a click away. Online shopping efficiency and consumption levels in three consumption domains. *Sustainable Production and Consumption*, *23*, 212–223.
<https://doi.org/10.1016/j.spc.2020.05.002>
- Frick, V., Matthies, E., Thøgersen, J., & Santarius, T. (2020). Do online environments promote sufficiency or overconsumption? Online advertisement and social media effects on clothing, digital devices, and air travel consumption. *Journal of Consumer Behaviour*, *45*(2), 205–220. <https://doi.org/10.1002/cb.1855>
- Gensler, S., Völckner, F., Egger, M., Fischbach, K., & Schoder, D. (2015). Listen to Your Customers: Insights into Brand Image Using Online Consumer-Generated Product Reviews. *International Journal of Electronic Commerce*, *20*(1), 112–141.
<https://doi.org/10.1080/10864415.2016.1061792>
- Gerber, A. S., Karlan, D., & Bergan, D. (2009). Does the Media Matter? A Field Experiment Measuring the Effect of Newspapers on Voting Behavior and Political Opinions. *American Economic Journal: Applied Economics*, *1*(2), 35–52.
<https://doi.org/10.1257/app.1.2.35>
- Ghanem, S. (1997). Filling in the tapestry: The second level of agenda setting. In M. E. McCombs, D. L. Shaw, & D. H. Weaver (Eds.), *Communication and democracy: Exploring the intellectual frontiers in agenda-setting theory* (pp. 3–14). Mahwah, NJ: Lawrence Erlbaum Associates.
- Golder, S. A., & Macy, M. W. (2014). Digital Footprints: Opportunities and Challenges for Online Social Research. *Annual Review of Sociology*, *40*(1), 129–152.
<https://doi.org/10.1146/annurev-soc-071913-043145>
- Gracia, A., & Magistris, T. de. (2008). The demand for organic foods in the South of Italy: A discrete choice model. *Food Policy*, *33*(5), 386–396.
<https://doi.org/10.1016/j.foodpol.2007.12.002>
- Grankvist, G., & Biel, A. (2001). The importance of beliefs and purchase criteria in the choice of eco-labelled food products. *Journal of Environmental Psychology*, *21*(4), 405–410. <https://doi.org/10.1006/jevps.2001.0234>
- Groshek, J., & Groshek, M. C. (2014). Agenda Trending: Reciprocity and the Predictive Capacity of Social Networking Sites in Intermedia Agenda Setting across Topics over Time. *Media and Communication*, *1*(1), 15–27.
<https://doi.org/10.17645/mac.v1i1.71>

- Grunert, K. G., Loose, S. M., Zhou, Y., & Tinggaard, S. (2015). Extrinsic and intrinsic quality cues in Chinese consumers' purchase of pork ribs. *Food Quality and Preference*, *42*, 37–47. <https://doi.org/10.1016/j.foodqual.2015.01.001>
- Guagnano, G. A., Stern, P. C., & Dietz, T. (1995). Influences on Attitude-Behavior Relationships. *Environment and Behavior*, *27*(5), 699–718. <https://doi.org/10.1177/0013916595275005>
- Gudrun Mahlau. (1999). *Das Image der Landwirtschaft*. Bonn: M. Wehle.
- Guo, L., Vargo, C. J., Pan, Z., Ding, W., & Ishwar, P. (2016). Big Social Data Analytics in Journalism and Mass Communication. *Journalism & Mass Communication Quarterly*, *93*(2), 332–359. <https://doi.org/10.1177/1077699016639231>
- Harper, G. C., & Makatouni, A. (2002). Consumer perception of organic food production and farm animal welfare. *British Food Journal*, *104*(3/4/5), 287–299. <https://doi.org/10.1108/00070700210425723>
- Harris, Z. S. (1954). Distributional Structure. *WORD*, *10*(2-3), 146–162. <https://doi.org/10.1080/00437956.1954.11659520>
- Hartmann, J., Huppertz, J., Schamp, C., & Heitmann, M. (2019). Comparing automated text classification methods. *International Journal of Research in Marketing*, *36*(1), 20–38. <https://doi.org/10.1016/j.ijresmar.2018.09.009>
- Hasselbach, J. L., & Roosen, J. (2015). Motivations behind preferences for local or organic food. *Journal of International Consumer Marketing*, *27*(4), 295–306. <https://doi.org/10.1080/08961530.2015.1022921>
- He, R., Lee, W. S., Ng, H. T., & Dahlmeier, D. (2017). An Unsupervised Neural Attention Model for Aspect Extraction. *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 388–397. <https://doi.org/10.18653/v1/P17-1036>
- Heise, H., & Theuvsen, L. (2017). Die Mehrzahlungsbereitschaft für Milch, Eier und Fleisch aus Tierwohlprogrammen: Eine repräsentative Verbraucherbefragung. *Journal of Consumer Protection and Food Safety*, *12*(2), 105–113. <https://doi.org/10.1007/s00003-016-1062-0>
- Hemmerling, S., Hamm, U., & Spiller, A. (2015). Consumption behaviour regarding organic food from a marketing perspective—a literature review. *Organic Agriculture*, *5*(4), 277–313. <https://doi.org/10.1007/s13165-015-0109-3>
- Hempel, C., & Hamm, U. (2016). Local and/or organic: A study on consumer preferences for organic food and food from different origins. *International Journal of Consumer Studies*, *40*(6), 732–741. <https://doi.org/10.1111/ijcs.12288>
- Hennig-Thurau, T., Gwinner, K. P., Walsh, G., & Gremler, D. D. (2004). Electronic word-of-mouth via consumer-opinion platforms: What motivates consumers to articulate themselves on the Internet? *Journal of Interactive Marketing*, *18*(1), 38–52. <https://doi.org/10.1002/dir.10073>
- Hester, J. B., & Gibson, R. (2003). The Economy and Second-Level Agenda Setting: A Time-Series Analysis of Economic News and Public Opinion about the Economy. *Journalism & Mass Communication Quarterly*, *80*(1), 73–90. <https://doi.org/10.1177/107769900308000106>

- Hjelmar, U. (2011). Consumers' purchase of organic food products. A matter of convenience and reflexive practices. *Appetite*, *56*(2), 336–344. <https://doi.org/10.1016/j.appet.2010.12.019>
- Hogan, B. (2010). The Presentation of Self in the Age of Social Media: Distinguishing Performances and Exhibitions Online. *Bulletin of Science, Technology & Society*, *30*(6), 377–386. <https://doi.org/10.1177/0270467610385893>
- Hossain, N., Krumm, J., & Gamon, M. (2019). "President Vows to Cut <Taxes> Hair": Dataset and Analysis of Creative Text Editing for Humorous Headlines. Retrieved from <https://arxiv.org/pdf/1906.00274>
- Hughner, R. S., McDonagh, P., Prothero, A., Shultz, C. J., & Stanton, J. (2007). Who are organic food consumers? A compilation and review of why people purchase organic food. *Journal of Consumer Behaviour*, *6*(2-3), 94–110. <https://doi.org/10.1002/cb.210>
- Humphreys, A., Wang, R. J.-H., Fischer, E., & Price, L. (2018). Automated Text Analysis for Consumer Research. *Journal of Consumer Research*, *44*(6), 1274–1306. <https://doi.org/10.1093/jcr/ucx104>
- IfD Allensbach. (2020). Internetnutzer in Deutschland nach Feldern der Internetnutzung im Jahr 2020. Retrieved from https://de-statista-com//statistik/daten/studie/940356/umfrage/umfrage-unter-internetnutzern-zu-den-beliebtesten-feldern-der-internetnutzung/*, (accessed September 28, 2020).
- Initiative D21 e.V. (2017). *Digital Index 2017/2018*. Berlin.
- Isoaho, K., Gritsenko, D., & Mäkelä, E. (2019). Topic Modeling and Text Analysis for Qualitative Policy Research. *Policy Studies Journal*, *49*(1), 300-325. <https://doi.org/10.1111/psj.12343>
- ITU. (2019). Measuring digital development. Retrieved from https://www.itu.int/en/ITU-D/Statistics/Documents/facts/FactsFigures2019.pdf*, (accessed December 3, 2020).
- IWV. (2020). Number of online and mobile visits on news portals in Germany in June 2020. Retrieved from https://www-statista-com/statistics/442988/most-visited-news-websites-germany/*, (accessed September 28, 2020).
- Jackson, P., Candel, J., Davies, A., Vries, H. de, Derani, C., Dragović-Uzelac, V., & ... (2020). *A sustainable food system for the European Union*. Berlin.
- Jacobi, C., van Atteveldt, W., & Welbers, K. (2015). Quantitative analysis of large amounts of journalistic texts using topic modelling. *Digital Journalism*, *4*(1), 89–106. <https://doi.org/10.1080/21670811.2015.1093271>
- Jacobsen, L. F., Tudoran, A. A., & Lähteenmäki, L. (2017). Consumers' motivation to interact in virtual food communities – The importance of self-presentation and learning. *Food Quality and Preference*, *62*, 8–16. <https://doi.org/10.1016/j.foodqual.2017.06.015>
- Janiszewski, C., & Wyer, R. S. (2014). Content and process priming: A review. *Journal of Consumer Psychology*, *24*(1), 96–118. <https://doi.org/10.1016/j.jcps.2013.05.006>

- Janssen, M. (2018). Determinants of organic food purchases: Evidence from household panel data. *Food Quality and Preference*, 68, 19–28. <https://doi.org/10.1016/j.foodqual.2018.02.002>
- Janssen, M., Heid, A., & Hamm, U. (2009). Is there a promising market ‘in between’ organic and conventional food? Analysis of consumer preferences. *Renewable Agriculture and Food Systems*, 24(03), 205–213. <https://doi.org/10.1017/S1742170509990056>
- Jensen, J. D., Christensen, T., Denver, S., Ditlevsen, K., Lassen, J., & Teuber, R. (2019). Heterogeneity in consumers' perceptions and demand for local (organic) food products. *Food Quality and Preference*, 73, 255–265. <https://doi.org/10.1016/j.foodqual.2018.11.002>
- Jonge, J. de, van der Lans, I. A., & van Trijp, H. C. M. (2015). Different shades of grey: Compromise products to encourage animal friendly consumption. *Food Quality and Preference*, 45, 87–99. <https://doi.org/10.1016/j.foodqual.2015.06.001>
- Juhl, H. J., Fenger, M. H. J., & Thøgersen, J. (2017). Will the Consistent Organic Food Consumer Step Forward? An Empirical Analysis. *Journal of Consumer Research*, 44(3), 519–535. <https://doi.org/10.1093/jcr/ucx052>
- Jungbluth, N., Tietje, O., & Scholz, R. W. (2000). Food purchases: Impacts from the consumers' point of view investigated with a modular LCA. *The International Journal of Life Cycle Assessment*, 5(3), 134–142. <https://doi.org/10.1007/BF02978609>
- Junsheng, H., Akhtar, R., Masud, M. M., Rana, M. S., & Banna, H. (2019). The role of mass media in communicating climate science: An empirical evidence. *Journal of Cleaner Production*, 238, 117934. <https://doi.org/10.1016/j.jclepro.2019.117934>
- Kahneman, D. (2011). *Thinking, Fast and Slow*. [S.I.]: Farrar, Straus and Giroux.
- Kassarjian, H. H. (1977). Content Analysis in Consumer Research. *Journal of Consumer Research*, 4(1), 8-18. <https://doi.org/10.1086/208674>
- Keller, K. L. (1991). Cue Compatibility and Framing in Advertising. *Journal of Marketing Research*, 28(1), 42–57. <https://doi.org/10.1177/002224379102800104>
- Kelley, H. H. (1967). Attribution theory in social psychology. *Nebraska Symposium on Motivation*. (15), 192–238.
- Kenny, D. A. (1975). Cross-lagged panel correlation: A test for spuriousness. *Psychological Bulletin*, 82(6), 887–903. <https://doi.org/10.1037/0033-2909.82.6.887>
- Kim, H. K., Kim, H., & Cho, S. (2017). Bag-of-concepts: Comprehending document representation through clustering words in distributed representation. *Neurocomputing*, 266, 336–352. <https://doi.org/10.1016/j.neucom.2017.05.046>
- Kitchen Stories. (2020a). Organic food purchasing behavior in Germany 2019: On average, how much of your shopping cart is filled with organic food? Retrieved from https://www.statista.com/statistics/1085285/organic-food-purchase-in-germany/*, (accessed October 15, 2020).
- Kitchen Stories. (2020b). Organic food purchasing behavior in the U.S. 2019: On average, how much of your shopping cart is filled with organic food? Retrieved from

- <https://www.statista.com/statistics/1085287/organic-food-purchase-in-the-us/>, (accessed October 15, 2020).
- Klöckner, C. A. (2013). A comprehensive model of the psychology of environmental behaviour—A meta-analysis. *Global Environmental Change*, 23(5), 1028–1038. <https://doi.org/10.1016/j.gloenvcha.2013.05.014>
- Koklic, M. K., Golob, U., Podnar, K., & Zabkar, V. (2019). The interplay of past consumption, attitudes and personal norms in organic food buying. *Appetite*, 137, 27–34. <https://doi.org/10.1016/j.appet.2019.02.010>
- Kolbe, R. H., & Burnett, M. S. (1991). Content-Analysis Research: An Examination of Applications with Directives for Improving Research Reliability and Objectivity. *Journal of Consumer Research*, 18(2), 243–250. <https://doi.org/10.1086/209256>
- Kozinets, R. V. (2002). The Field behind the Screen: Using Netnography for Marketing Research in Online Communities. *Journal of Marketing Research*, 39(1), 61–72. <https://doi.org/10.1509/jmkr.39.1.61.18935>
- Krippendorff, K. (2019). *Content analysis: An introduction to its methodology* (4th ed.). Los Angeles, London, New Delhi, Singapore: SAGE.
- Ksiazek, T. B. (2018). Commenting on the News. *Journalism Studies*, 19(5), 650–673. <https://doi.org/10.1080/1461670X.2016.1209977>
- Ksiazek, T. B., Peer, L., & Lessard, K. (2016). User engagement with online news: Conceptualizing interactivity and exploring the relationship between online news videos and user comments. *New Media & Society*, 18(3), 502–520. <https://doi.org/10.1177/1461444814545073>
- Kushwah, S., Dhir, A., Sagar, M., & Gupta, B. (2019). Determinants of organic food consumption. A systematic literature review on motives and barriers. *Appetite*, 143, 104402. <https://doi.org/10.1016/j.appet.2019.104402>
- Lancaster, K. J. (1966). A New Approach to Consumer Theory. *Journal of Political Economy*, 74(2), 132–157. <https://doi.org/10.1086/259131>
- Le, Q. V., & Mikolov, T. Distributed representations of sentences and documents. In *International Conference on Machine Learning* (pp. 1188–1196).
- Lee, E.-J. (2012). That's Not the Way It Is: How User-Generated Comments on the News Affect Perceived Media Bias. *Journal of Computer-Mediated Communication*, 18(1), 32–45. <https://doi.org/10.1111/j.1083-6101.2012.01597.x>
- Lee, E.-J., Kim, H. S., & Cho, J. (2017). How user comments affect news processing and reality perception: Activation and refutation of regional prejudice. *Communication Monographs*, 84(1), 75–93. <https://doi.org/10.1080/03637751.2016.1231334>
- Lee, E.-J., & Tandoc, E. C. (2017). When News Meets the Audience: How Audience Feedback Online Affects News Production and Consumption. *Human Communication Research*, 43(4), 436–449. <https://doi.org/10.1111/hcre.12123>
- Lee, H.-J., & Hwang, J. (2016). The driving role of consumers' perceived credence attributes in organic food purchase decisions: A comparison of two groups of consumers. *Food Quality and Preference*, 54, 141–151. <https://doi.org/10.1016/j.foodqual.2016.07.011>

- Lee, H.-J., & Yun, Z.-S. (2015). Consumers' perceptions of organic food attributes and cognitive and affective attitudes as determinants of their purchase intentions toward organic food. *Food Quality and Preference*, *39*, 259–267. <https://doi.org/10.1016/j.foodqual.2014.06.002>
- Lim, J. (2006). A Cross-Lagged Analysis of Agenda Setting among Online News Media. *Journalism & Mass Communication Quarterly*, *83*(2), 298–312. <https://doi.org/10.1177/107769900608300205>
- Lim, J. (2011). First-level and second-level intermedia agenda-setting among major news websites. *Asian Journal of Communication*, *21*(2), 167–185. <https://doi.org/10.1080/01292986.2010.539300>
- Lockie, S. (2006). Capturing the Sustainability Agenda: Organic Foods and Media Discourses on Food Scares, Environment, Genetic Engineering, and Health. *Agriculture and Human Values*, *23*(3), 313–323. <https://doi.org/10.1007/s10460-006-9007-3>
- Loewenstein, G. (2000). Emotions in Economic Theory and Economic Behavior. *American Economic Review*, *90*(2), 426–432. <https://doi.org/10.1257/aer.90.2.426>
- Lombard, M., Snyder-Duch, J., & Bracken, C. C. (2002). Content analysis in mass communication: Assessment and reporting of intercoder reliability. *Human Communication Research*, *28*(4), 587–604. <https://doi.org/10.1111/j.1468-2958.2002.tb00826.x>
- Louviere, J. J., Hensher, D. A., Swait, J., & Adamowicz, W. L. (2010). *Stated choice methods: Analysis and applications* (7. printing). Cambridge: Cambridge Univ. Press.
- Lusk, J., & Norwood, F. B. (2009). Bridging the gap between laboratory experiments and naturally occurring markets: An inferred valuation method. *Journal of Environmental Economics and Management*, *58*(2), 236–250. <https://doi.org/10.1016/j.jeem.2008.12.003>
- Lusk, J., Schroeder, T. C., & Tonsor, G. T. (2014). Distinguishing beliefs from preferences in food choice. *European Review of Agricultural Economics*, *41*(4), 627–655. <https://doi.org/10.1093/erae/jbt035>
- Lusk, J. L. (2011). External validity of the food values scale. *Food Quality and Preference*, *22*(5), 452–462. <https://doi.org/10.1016/j.foodqual.2011.02.009>
- Lyu, F., & Choi, J. (2020). The Forecasting Sales Volume and Satisfaction of Organic Products through Text Mining on Web Customer Reviews. *Sustainability*, *12*(11), 4383. <https://doi.org/10.3390/su12114383>
- Maison, D., Greenwald, A. G., & Bruin, R. H. (2004). Predictive Validity of the Implicit Association Test in Studies of Brands, Consumer Attitudes, and Behavior. *Journal of Consumer Psychology*, *14*(4), 405–415. https://doi.org/10.1207/s15327663jcp1404_9
- Mangold, W. G., & Faulds, D. J. (2009). Social media: The new hybrid element of the promotion mix. *Business Horizons*, *52*(4), 357–365. <https://doi.org/10.1016/j.bushor.2009.03.002>

- Manning, C. D., & Schütze, H. (2005). *Foundations of statistical natural language processing* (8th ed.). Cambridge, Mass.: MIT Press.
- Manski, C. F., & McFadden, D. T. (1981). Econometric models of probabilistic choice. In C. F. Manski & D. T. McFadden (Eds.), *Structural analysis of discrete data with econometric applications* (pp. 198–272). Cambridge, Mass.: MIT Press.
- Markham, A. N., Tiidenberg, K., & Herman, A. (2018). Ethics as Methods: Doing Ethics in the Era of Big Data Research—Introduction. *Social Media + Society*, 4(3), 205630511878450. <https://doi.org/10.1177/2056305118784502>
- Massey, M., O'Cass, A., & Otahal, P. (2018). A meta-analytic study of the factors driving the purchase of organic food. *Appetite*, 125, 418–427. <https://doi.org/10.1016/j.appet.2018.02.029>
- McCombs, M. E. (2014). *Setting the agenda: The mass media and public opinion* (2nd ed.). Cambridge: Polity Press.
- McCombs, M. E., & Shaw, D. L. (1972). The Agenda-Setting Function of Mass Media. *Public Opinion Quarterly*, 36(2), 176–187. <https://doi.org/10.1086/267990>
- McFadden, D. T. (1974). Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed.), *Frontiers in Econometrics* (pp. 105–142). New York: Academic Press.
- Meyers, C., & Abrams, K. M. (2010). Feeding the debate: a qualitative framing analysis of organic food news media coverage. *Journal of Applied Communications*, 94(3-4), 22–37.
- Meza, X. V., & Park, H. W. (2016). Organic products in Mexico and South Korea on Twitter. *Journal of Business Ethics*, 135(3), 587–603. <https://doi.org/10.1007/s10551-014-2345-y>
- Midden, C. J. H., Kaiser, F. G., & Teddy McCalley, L. (2007). Technology's Four Roles in Understanding Individuals' Conservation of Natural Resources. *Journal of Social Issues*, 63(1), 155–174. <https://doi.org/10.1111/j.1540-4560.2007.00501.x>
- Mie, A., Andersen, H. R., Gunnarsson, S., Kahl, J., Kesse-Guyot, E., Rembialkowska, E., . . . Grandjean, P. (2017). Human health implications of organic food and organic agriculture: a comprehensive review. *Environmental Health*, 16(1), 111. <https://doi.org/10.1186/s12940-017-0315-4>
- Mikolov, T., Chen, K., Corrado, G., & Dean, J. Efficient Estimation of Word Representations in Vector Space. *ArXiv*. Retrieved from <http://arxiv.org/pdf/1301.3781v3>
- Minton, E. A., Cornwell, T. B., & Kahle, L. R. (2017). A theoretical review of consumer priming: Prospective theory, retrospective theory, and the affective-behavioral-cognitive model. *Journal of Consumer Behaviour*, 16(4), 309–321. <https://doi.org/10.1002/cb.1624>
- Moser, A. K. (2016). Buying organic – decision-making heuristics and empirical evidence from Germany. *Journal of Consumer Marketing*, 33(7), 552–561. <https://doi.org/10.1108/JCM-04-2016-1790>

- Mudambi, & Schuff. (2010). Research Note: What Makes a Helpful Online Review? A Study of Customer Reviews on Amazon.com. *MIS Quarterly*, 34(1), 185-200. <https://doi.org/10.2307/20721420>
- Müller, A., Schader, C., El-Hage Scialabba, N., Brüggemann, J., Isensee, A., Erb, K.-H., . . . Niggli, U. (2017). Strategies for feeding the world more sustainably with organic agriculture. *Nature Communications*, 8(1), 1290. <https://doi.org/10.1038/s41467-017-01410-w>
- Müller, C. E., & Gaus, H. (2015). Consumer Response to Negative Media Information About Certified Organic Food Products. *Journal of Consumer Policy*, 38(4), 387–409. <https://doi.org/10.1007/s10603-015-9299-z>
- Napoli, J., Dickinson, S. J., Beverland, M. B., & Farrelly, F. (2014). Measuring consumer-based brand authenticity. *Journal of Business Research*, 67(6), 1090–1098. <https://doi.org/10.1016/j.jbusres.2013.06.001>
- Nayakankuppam, D., Priester, J. R., Kwon, J. H., Donovan, L. A. N., & Petty, R. E. (2018). Construction and retrieval of evaluative judgments: The attitude strength moderation model. *Journal of Experimental Social Psychology*, 76, 54–66. <https://doi.org/10.1016/j.jesp.2017.12.005>
- Nelson, D. L., Bennett, D. J., Gee, N. R., Schreiber, T. A., & McKinney, V. M. (1993). Implicit memory: Effects of network size and interconnectivity on cued recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19(4), 747–764. <https://doi.org/10.1037/0278-7393.19.4.747>
- Nelson, P. (1970). Information and Consumer Behavior. *Journal of Political Economy*. (78), 311–329.
- Netzer, O., Feldman, R., Goldenberg, J., & Fresko, M. (2012). Mine Your Own Business: Market-Structure Surveillance Through Text Mining. *Marketing Science*, 31(3), 521–543. <https://doi.org/10.1287/mksc.1120.0713>
- Neuman, W. R., Guggenheim, L., Mo Jang, S., & Bae, S. Y. (2014). The Dynamics of Public Attention: Agenda-Setting Theory Meets Big Data. *Journal of Communication*, 64(2), 193–214. <https://doi.org/10.1111/jcom.12088>
- NTIA. (2018). Digital Nation Data Explorer. Retrieved from https://www.ntia.doc.gov/data/digital-nation-data-explorer*, (accessed 14 January 2019).
- Nuttavuthisit, K., & Thøgersen, J. (2017). The importance of consumer trust for the emergence of a market for green products: The Case of Organic Food. *Journal of Business Ethics*, 140(2), 323–337. <https://doi.org/10.1007/s10551-015-2690-5>
- Olson, E. L. (2017). The rationalization and persistence of organic food beliefs in the face of contrary evidence. *Journal of Cleaner Production*, 140, 1007–1013. <https://doi.org/10.1016/j.jclepro.2016.06.005>
- Ordenes, F. V., Ludwig, S., Ruyter, K. de, Grewal, D., & Wetzels, M. (2017). Unveiling What Is Written in the Stars: Analyzing Explicit, Implicit and Discourse Patterns of Sentiment in Social Media. *Journal of Consumer Research*, 74(6), 875–894. <https://doi.org/10.1093/jcr/ucw070>

- Organic Food Production Alliance (BÖLW). (n.d.). Wie ist die Ökologische Landwirtschaft entstanden? Retrieved from <https://www.boelw.de/service/bio-faq/bio-basics/artikel/wie-ist-die-oekologische-landwirtschaft-entstanden/>*, (accessed November 12, 2020).
- Organic Food Production Alliance (BÖLW). (2018). *Die Bio-Branche 2017*. Berlin.
- Organic Food Production Alliance (BÖLW). (2020). *Branchen Report 2020*. Berlin.
- OTA. (2018). Market Analysis. Retrieved from www.ota.com/
- Padel, S., & Foster, C. (2005). Exploring the gap between attitudes and behaviour: Understanding why consumers buy or do not buy organic food. *British Food Journal*, 107(8), 606–625. <https://doi.org/10.1108/00070700510611002>
- Pearson, D., Henryks, J., & Jones, H. (2010). Organic food: What we know (and do not know) about consumers. *Renewable Agriculture and Food Systems*, 26(02), 171–177. <https://doi.org/10.1017/S1742170510000499>
- Pedersen, S., Aschemann-Witzel, J., & Thøgersen, J. (2018). Consumers' evaluation of imported organic food products: The role of geographical distance. *Appetite*, 130, 134–145. <https://doi.org/10.1016/j.appet.2018.08.016>
- Peschel, A. O., Kazemi, S., Liebichová, M., Sarraf, S. C. M., & Aschemann-Witzel, J. (2019). Consumers associative networks of plant-based food product communications. *Food Quality and Preference*, 75, 145-156. <https://doi.org/10.1016/j.foodqual.2019.02.015>
- Peter, J. P., & Olson, J. C. (2010). *Consumer Behavior and Marketing Strategy* (9th ed.). New York: McGraw-Hill.
- Petty, R. E., Fazio, R. H., & Briñol, P. (2009). The New Implicit Measures: An Overview. In R. E. Petty, R. H. Fazio, & P. Briñol Turnes (Eds.), *Occupational Safety & Health Guide Series. Attitudes: Insights from the new implicit measures* (pp. 3–18). New York: Routledge.
- Phillips, K. A., Johnson, F. R., & Maddala, T. (2002). Measuring what people value: a comparison of "attitude" and "preference" surveys. *Health Services Research*, 37(6), 1659–1679. <https://doi.org/10.1111/1475-6773.01116>
- Pinto, S., Albanese, F., Dorso, C. O., & Balenzuela, P. (2019). Quantifying time-dependent Media Agenda and public opinion by topic modeling. *Physica a: Statistical Mechanics and Its Applications*, 524, 614–624. <https://doi.org/10.1016/j.physa.2019.04.108>
- Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *The Journal of Applied Psychology*, 88(5), 879–903. <https://doi.org/10.1037/0021-9010.88.5.879>
- Pollach, I. (2012). Taming Textual Data: The Contribution of Corpus Linguistics to Computer-Aided Text Analysis. *Organizational Research Methods*, 15(2), 263–287. <https://doi.org/10.1177/1094428111417451>
- Pollach, I. (2014). Corporate Environmental Reporting and News Coverage of Environmental Issues: an Agenda-Setting Perspective. *Business Strategy and the Environment*, 23(5), 349–360. <https://doi.org/10.1002/bse.1792>

- Pomerantz, E. M., Chaiken, S., & Tordesillas, R. S. (1995). Attitude strength and resistance processes. *Journal of Personality and Social Psychology*, *69*(3), 408–419. <https://doi.org/10.1037/0022-3514.69.3.408>
- Poore, J., & Nemecek, T. (2018). Reducing food's environmental impacts through producers and consumers. *Science (New York, N.Y.)*, *360*(6392), 987–992. <https://doi.org/10.1126/science.aaq0216>
- Posavac, S. S., Sanbonmatsu, D. M., & Fazio, R. H. (1997). Considering the best choice: Effects of the salience and accessibility of alternatives on attitude–decision consistency. *Journal of Personality and Social Psychology*, *72*(2), 253–261. <https://doi.org/10.1037/0022-3514.72.2.253>
- Prothero, A., Dobscha, S., Freund, J., Kilbourne, W. E., Luchs, M. G., Ozanne, L. K., & Thøgersen, J. (2011). Sustainable Consumption: Opportunities for Consumer Research and Public Policy. *Journal of Public Policy & Marketing*, *30*(1), 31–38. <https://doi.org/10.1509/jppm.30.1.31>
- Rana, J., & Paul, J. (2017). Consumer behavior and purchase intention for organic food: A review and research agenda. *Journal of Retailing and Consumer Services*, *38*, 157–165. <https://doi.org/10.1016/j.jretconser.2017.06.004>
- Rana, J., & Paul, J. (2020). Health motive and the purchase of organic food: A meta-analytic review. *International Journal of Consumer Studies*, *44*(2), 162–171. <https://doi.org/10.1111/ijcs.12556>
- Randall, D. M., & Fernandes, M. F. (1991). The social desirability response bias in ethics research. *Journal of Business Ethics*, *10*(11), 805–817. <https://doi.org/10.1007/BF00383696>
- Reisch, L., Eberle, U., & Lorek, S. (2013). Sustainable food consumption: an overview of contemporary issues and policies. *Sustainability: Science, Practice and Policy*, *9*(2), 7–25. <https://doi.org/10.1080/15487733.2013.11908111>
- Reisch, L. A. (2001). The Internet and Sustainable Consumption: Perspectives on a Janus Face. *Journal of Consumer Policy*, *24*(3-4), 251–286. <https://doi.org/10.1023/A:1013977509623>
- Reisch, L. A., & Zhao, M. (2017). Behavioural economics, consumer behaviour and consumer policy: state of the art. *Behavioural Public Policy*, *1*(2), 190–206. <https://doi.org/10.1017/bpp.2017.1>
- Renz, S. M., Carrington, J. M., & Badger, T. A. (2018). Two Strategies for Qualitative Content Analysis: An Intramethod Approach to Triangulation. *Qualitative Health Research*, *28*(5), 824–831. <https://doi.org/10.1177/1049732317753586>
- Rogosa, D. (1980). A critique of cross-lagged correlation. *Psychological Bulletin*, *88*(2), 245–258. <https://doi.org/10.1037/0033-2909.88.2.245>
- Rozelle, R. M., & Campbell, D. T. (1969). More plausible rival hypotheses in the cross-lagged panel correlation technique. *Psychological Bulletin*, *71*(1), 74–80. <https://doi.org/10.1037/h0026863>
- Ruggeri, A., & Samoggia, A. (2018). Twitter communication of agri-food chain actors on palm oil environmental, socio-economic, and health sustainability. *Journal of Consumer Behaviour*, *17*(1), 75–93. <https://doi.org/10.1002/cb.1699>

- Sahota, A. (2018). The global market for organic food & drink. In IFOAM & FiBL - Organics International (Ed.), *The World of Organic Agriculture: Statistics and Emerging Trends 2018* (pp. 145–149).
- Santana, A. D. (2011). Online Readers' Comments Represent New Opinion Pipeline. *Newspaper Research Journal*, 32(3), 66–81.
<https://doi.org/10.1177/073953291103200306>
- Scalco, A., Noventa, S., Sartori, R., & Ceschi, A. (2017). Predicting organic food consumption: A meta-analytic structural equation model based on the theory of planned behavior. *Appetite*, 112, 235–248.
<https://doi.org/10.1016/j.appet.2017.02.007>
- Schäufele, I., & Hamm, U. (2018). Organic wine purchase behaviour in Germany: Exploring the attitude-behaviour-gap with data from a household panel. *Food Quality and Preference*, 63, 1–11. <https://doi.org/10.1016/j.foodqual.2017.07.010>
- Schleenbecker, R., & Hamm, U. (2013). Consumers' perception of organic product characteristics. A review. *Appetite*, 71, 420–429.
<https://doi.org/10.1016/j.appet.2013.08.020>
- Schmitt, M., & Schuller, B. (2017). OpenXBOW: introducing the passau open-source crossmodal bag-of-words toolkit. *The Journal of Machine Learning Research*, 18(1).
- Seufert, V., Ramankutty, N., & Foley, J. A. (2012). Comparing the yields of organic and conventional agriculture. *Nature*, 485(7397), 229–232.
<https://doi.org/10.1038/nature11069>
- SevenOne Media. (2020). Media Activity Guide 2020: Average daily usage time of selected types of media in Germany in 2019. Retrieved from https://www.statista.com/statistics/380121/media-daily-usage-time-germany/*, (accessed October 19, 2020).
- Shahin, S. (2016). When Scale Meets Depth: Integrating Natural Language Processing and Textual Analysis for Studying Digital Corpora. *Communication Methods and Measures*, 10(1), 28–50. <https://doi.org/10.1080/19312458.2015.1118447>
- Shavitt, S., & Fazio, R. H. (1991). Effects of Attribute Salience on the Consistency between Attitudes and Behavior Predictions. *Personality and Social Psychology Bulletin*, 17(5), 507–516. <https://doi.org/10.1177/0146167291175005>
- Siegrist, M., Visschers, V. H., & Hartmann, C. (2015). Factors influencing changes in sustainability perception of various food behaviors: Results of a longitudinal study. *Food Quality and Preference*, 46, 33–39.
<https://doi.org/10.1016/j.foodqual.2015.07.006>
- Sloan, L., & Quan-Haase, A. (2017). *The SAGE Handbook of Social Media Research Methods*. London: SAGE Publications Ltd.
- Smith-Spangler, C., Brandeau, M. L., Hunter, G. E., Bavinger, J. C., Pearson, M., Eschbach, P. J., . . . Bravata, D. M. (2012). Are organic foods safer or healthier than conventional alternatives? A systematic review. *Annals of Internal Medicine*, 157(5), 348–366. <https://doi.org/10.7326/0003-4819-157-5-201209040-00007>

- Spoolder, H. am. (2007). Animal welfare in organic farming systems. *Journal of the Science of Food and Agriculture*, 87(15), 2741–2746.
<https://doi.org/10.1002/jsfa.2999>
- Sridhar, R., & Kumar, V. (2015). Unsupervised Topic Modeling for Short Texts Using Distributed Representations of Words. In *Proceedings of the 1st Workshop on Vector Space Modeling for Natural Language Processing* (pp. 192–200).
- Steffen, W., Richardson, K., Rockström, J., Cornell, S. E., Fetzer, I., Bennett, E. M., . . . Sörlin, S. (2015). Sustainability. Planetary boundaries: guiding human development on a changing planet. *Science*, 347(6223), 736–747.
<https://doi.org/10.1126/science.1259855>
- Steinert-Threkeld, Z. C. (2018). *Twitter as Data*. Cambridge: Cambridge University Press.
- Stern, P. C. (2000). New Environmental Theories: Toward a Coherent Theory of Environmentally Significant Behavior. *Journal of Social Issues*, 56(3), 407–424.
<https://doi.org/10.1111/0022-4537.00175>
- Stern, P. C., Dietz, T., Abel, T., Guagnano, G. A., & Kalof, L. (1999). A Value-Belief-Norm Theory of Support for Social Movements: The Case of Environmentalism. *Human Ecology Review*, 6(2), 81–97.
- Street, D. J., Burgess, L., & Louviere, J. J. (2005). Quick and easy choice sets: Constructing optimal and nearly optimal stated choice experiments. *International Journal of Research in Marketing*, 22(4), 459–470.
<https://doi.org/10.1016/j.ijresmar.2005.09.003>
- Swait, J., & Adamowicz, W. (2001). The Influence of Task Complexity on Consumer Choice: A Latent Class Model of Decision Strategy Switching. *Journal of Consumer Research*, 28(1), 135–148. <https://doi.org/10.1086/321952>
- Symoneaux, R., Galmarini, M. V., & Mehinagic, E. (2012). Comment analysis of consumer's likes and dislikes as an alternative tool to preference mapping. A case study on apples. *Food Quality and Preference*, 24(1), 59–66.
<https://doi.org/10.1016/j.foodqual.2011.08.013>
- Takeshita, T. (2006). Current Critical Problems in Agenda-Setting Research. *International Journal of Public Opinion Research*, 18(3), 275–296.
<https://doi.org/10.1093/ijpor/edh104>
- Tang, T., Fang, E., & Wang, F. (2014). Is Neutral Really Neutral? The Effects of Neutral User-Generated Content on Product Sales. *Journal of Marketing*, 78(4), 41–58. <https://doi.org/10.1509/jm.13.0301>
- Tausczik, Y. R., & Pennebaker, J. W. (2010). The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods. *Journal of Language and Social Psychology*, 29(1), 24–54. <https://doi.org/10.1177/0261927X09351676>
- Testa, F., Sarti, S., & Frey, M. (2019). Are green consumers really green? Exploring the factors behind the actual consumption of organic food products. *Business Strategy and the Environment*, 28(2), 327–338. <https://doi.org/10.1002/bse.2234>
- Thaler, R. H., & Sunstein, C. R. (2009). *Nudge: Improving decisions about health, wealth, and happiness*. New York, NY: Penguin.

- Thøgersen, J. (2006). Media attention and the market for 'green' consumer products. *Business Strategy and the Environment*, *15*(3), 145–156. <https://doi.org/10.1002/bse.521>
- Thøgersen, J. (2010). Country differences in sustainable consumption: The case of organic food. *Journal of Macromarketing*, *30*(2), 171–185. <https://doi.org/10.1177/0276146710361926>
- Thøgersen, J. (2014). Unsustainable Consumption. *European Psychologist*, *19*(2), 84–95. <https://doi.org/10.1027/1016-9040/a000176>
- Thøgersen, J., & Alfinito, S. (2020). Goal activation for sustainable consumer choices: A comparative study of Denmark and Brazil. *Journal of Consumer Behaviour*, *19*(6), 556–569. <https://doi.org/10.1002/cb.1824>
- Thøgersen, J., Barcellos, M. D. de, Perin, M. G., & Zhou, Y. (2015). Consumer buying motives and attitudes towards organic food in two emerging markets. *International Marketing Review*, *32*(3/4), 389–413. <https://doi.org/10.1108/IMR-06-2013-0123>
- Thøgersen, J., & Ölander, F. (2006). To What Degree are Environmentally Beneficial Choices Reflective of a General Conservation Stance? *Environment and Behavior*, *38*(4), 550–569. <https://doi.org/10.1177/0013916505283832>
- Thøgersen, J., Pedersen, S., & Aschemann-Witzel, J. (2019). The impact of organic certification and country of origin on consumer food choice in developed and emerging economies. *Food Quality and Preference*, *72*, 10–30. <https://doi.org/10.1016/j.foodqual.2018.09.003>
- Thøgersen, J., Pedersen, S., Paternoga, M., Schwendel, E., & Aschemann-Witzel, J. (2017). How important is country-of-origin for organic food consumers? A review of the literature and suggestions for future research. *British Food Journal*, *119*(3), 542–557. <https://doi.org/10.1108/BFJ-09-2016-0406>
- Thorsøe, M. H. (2015). Maintaining trust and credibility in a continuously evolving organic food system. *Journal of Agricultural and Environmental Ethics*, *28*(4), 767–787. <https://doi.org/10.1007/s10806-015-9559-6>
- Thurmond, V. A. (2001). The point of triangulation. *Journal of Nursing Scholarship*, *33*(3), 253–258. <https://doi.org/10.1111/j.1547-5069.2001.00253.x>
- Timoshenko, A., & Hauser, J. R. (2019). Identifying Customer Needs from User-Generated Content. *Marketing Science*, *38*(1), 1–20. <https://doi.org/10.1287/mksc.2018.1123>
- Tirunillai, S., & Tellis, G. J. (2012). Does Chatter Really Matter? Dynamics of User-Generated Content and Stock Performance. *Marketing Science*, *31*(2), 198–215. <https://doi.org/10.1287/mksc.1110.0682>
- Tirunillai, S., & Tellis, G. J. (2014). Mining Marketing Meaning from Online Chatter: Strategic Brand Analysis of Big Data Using Latent Dirichlet Allocation. *Journal of Marketing Research*, *51*(4), 463–479. <https://doi.org/10.1509/jmr.12.0106>
- Tobler, C., Visschers, V. H. M., & Siegrist, M. (2011). Eating green. Consumers' willingness to adopt ecological food consumption behaviors. *Appetite*, *57*(3), 674–682. <https://doi.org/10.1016/j.appet.2011.08.010>

- Tonsor, G. T., Wolf, C., & Olynk, N. (2009). Consumer voting and demand behavior regarding swine gestation crates. *Food Policy*, *34*(6), 492–498.
<https://doi.org/10.1016/j.foodpol.2009.06.008>
- Trivedi, R. H., Patel, J. D., & Acharya, N. (2018). Causality analysis of media influence on environmental attitude, intention and behaviors leading to green purchasing. *Journal of Cleaner Production*, *196*, 11–22.
<https://doi.org/10.1016/j.jclepro.2018.06.024>
- Tsiafouli, M. A., Thébault, E., Sgardelis, S. P., Ruiter, P. C. de, van der Putten, W. H., Birkhofer, K., . . . Hedlund, K. (2015). Intensive agriculture reduces soil biodiversity across Europe. *Global Change Biology*, *21*(2), 973–985.
<https://doi.org/10.1111/gcb.12752>
- Tuck, S. L., Winqvist, C., Mota, F., Ahnström, J., Turnbull, L. A., & Bengtsson, J. (2014). Land-use intensity and the effects of organic farming on biodiversity: a hierarchical meta-analysis. *The Journal of Applied Ecology*, *51*(3), 746–755.
<https://doi.org/10.1111/1365-2664.12219>
- Tukker, A., Cohen, M. J., Hubacek, K., & Mont, O. (2010). The Impacts of Household Consumption and Options for Change. *Journal of Industrial Ecology*, *14*(1), 13–30.
<https://doi.org/10.1111/j.1530-9290.2009.00208.x>
- Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science (New York, N.Y.)*, *185*(4157), 1124–1131.
<https://doi.org/10.1126/science.185.4157.1124>
- United Nations. (2020). The 17 goals. Retrieved from https://sdgs.un.org/goals*, (accessed October 20, 2020).
- USDA Agricultural Marketing Service. (2016). Labeling Organic Products. Retrieved from https://www.ams.usda.gov/publications/content/labeling-organic-products*, (accessed October 19, 2020).
- USDA Economic Research Service. (2020). Organic Agriculture. Retrieved from https://www.ers.usda.gov/topics/natural-resources-environment/organic-agriculture/*, (accessed October 19, 2020).
- Van Ittersum, K., Pennings, J. M., Wansink, B., & van Trijp, H. C. (2007). The validity of attribute-importance measurement: A review. *Journal of Business Research*, *60*(11), 1177–1190. <https://doi.org/10.1016/j.jbusres.2007.04.001>
- Vargo, C. J., Guo, L., McCombs, M., & Shaw, D. L. (2014). Network Issue Agendas on Twitter During the 2012 U.S. Presidential Election. *Journal of Communication*, *64*(2), 296–316. <https://doi.org/10.1111/jcom.12089>
- Vega-Zamora, M., Parras-Rosa, M., Murgado-Armenteros, E., & Torres-Ruiz, F. J. (2013). The influence of the term ‘organic’ on organic food purchasing behavior. *Procedia - Social and Behavioral Sciences*, *81*, 660–671.
<https://doi.org/10.1016/j.sbspro.2013.06.493>
- Vega-Zamora, M., Torres-Ruiz, F. J., Murgado-Armenteros, E. M., & Parras-Rosa, M. (2014). Organic as a Heuristic Cue: What Spanish Consumers Mean by Organic Foods. *Psychology and Marketing*, *31*(5), 349–359.
<https://doi.org/10.1002/mar.20699>

- Verain, M. C. D., Dagevos, H., & Antonides, G. (2015). Sustainable food consumption. Product choice or curtailment? *Appetite*, *91*, 375–384. <https://doi.org/10.1016/j.appet.2015.04.055>
- Vermeir, I., & Verbeke, W. (2006). Sustainable Food Consumption: Exploring the Consumer “Attitude – Behavioral Intention” Gap. *Journal of Agricultural and Environmental Ethics*, *19*(2), 169–194. <https://doi.org/10.1007/s10806-005-5485-3>
- Vermunt, J. K., & Magidson, J. (2016). *Upgrade Manual for Latent Gold 5.1*.
- Vidal, L., Ares, G., & Jaeger, S. R. (2018). Chapter 6 - Application of Social Media for Consumer Research. In G. Ares & P. Varela (Eds.), *Woodhead Publishing Series in Food Science, Technology and Nutrition. Methods in consumer research* (pp. 125–155). Duxford, United Kingdom: Woodhead Publishing.
- Vidal, L., Ares, G., Machín, L., & Jaeger, S. R. (2015). Using Twitter data for food-related consumer research: A case study on “what people say when tweeting about different eating situations”. *Food Quality and Preference*, *45*, 58–69. <https://doi.org/10.1016/j.foodqual.2015.05.006>
- Visalli, M., Mahieu, B., Thomas, A., & Schlich, P. (2020). Automated sentiment analysis of Free-Comment: An indirect liking measurement? *Food Quality and Preference*, *82*, 103888. <https://doi.org/10.1016/j.foodqual.2020.103888>
- Vittersø, G., & Tangeland, T. (2015). The role of consumers in transitions towards sustainable food consumption. The case of organic food in Norway. *Journal of Cleaner Production*, *92*, 91–99. <https://doi.org/10.1016/j.jclepro.2014.12.055>
- VuMA. (2018). *VuMA Berichtsband 2018*. Frankfurt am Main, Hamburg, Mainz.
- Wanta, W., King, P., & McCombs, M. E. (1995). A comparison of factors influencing issue diversity in the U.S. and Taiwan. *International Journal of Public Opinion Research*, *7*(4), 353–365. <https://doi.org/10.1093/ijpor/7.4.353>
- Weaver, D. H. (2007). Thoughts on Agenda Setting, Framing, and Priming. *Journal of Communication*, *57*(1), 142–147. <https://doi.org/10.1111/j.1460-2466.2006.00333.x>
- Westaby, J. D. (2005). Behavioral reasoning theory: Identifying new linkages underlying intentions and behavior. *Organizational Behavior and Human Decision Processes*, *98*(2), 97–120. <https://doi.org/10.1016/j.obhdp.2005.07.003>
- Willer, H., Sampson, G., Voora, V., Dang, D., & Lernoud, J. (2019). *The State of Sustainable Markets 2019: Statistics and Emerging Trends*. Geneva: International Trade Centre.
- Willer, H., Schlatter, B., Trávníček, J., Kemper, L., & Lernoud, J. (Eds.). (2020). *The World of Organic Agriculture: Statistics and Emerging Trends 2020*. Frick, Switzerland.
- Wilson, A. L., Buckley, E., Buckley, J. D., & Bogomolova, S. (2016). Nudging healthier food and beverage choices through salience and priming. Evidence from a systematic review. *Food Quality and Preference*, *51*, 47–64. <https://doi.org/10.1016/j.foodqual.2016.02.009>
- Winter, S., Brückner, C., & Krämer, N. C. (2015). They Came, They Liked, They Commented: Social Influence on Facebook News Channels. *Cyberpsychology*,

- Behavior and Social Networking*, 18(8), 431–436.
<https://doi.org/10.1089/cyber.2015.0005>
- World Bank. (2003). *World development report 2003: Sustainable development in a dynamic economy*. Washington, D.C., New York: The World Bank; Oxford University Press.
- Xie, J., Gao, Z., Swisher, M., & Zhao, X. (2016). Consumers' preferences for fresh broccolis: Interactive effects between country of origin and organic labels. *Agricultural Economics*, 47(2), 181–191. <https://doi.org/10.1111/agec.12193>
- Xing, C., Wang, D., Zhang, X., & Liu, C. (2014). Document classification with distributions of word vectors. In *Signal and Information Processing Association Annual Summit and Conference (APSIPA), 2014 Asia-Pacific*. IEEE.
<https://doi.org/10.1109/APSIPA.2014.7041633>
- Yadavalli, A., & Jones, K. (2014). Does media influence consumer demand? The case of lean finely textured beef in the United States. *Food Policy*, 49, 219–227.
<https://doi.org/10.1016/j.foodpol.2014.08.002>
- Yang, Y., Abrego, G. H., Yuan, S., Guo, M., Shen, Q., Cer, D., . . . Kurzweil, R. (2018). English and German language-agnostic text encoder. Retrieved from https://tfhub.dev/google/universal-sentence-encoder-xling/en-de/1*, (accessed June 15, 2020).
- Yang, Z., Dai, Z., Yang, Y., Carbonell, J., Salakhutdinov, R. R., & Le, Q. V. (2019). XLNet: Generalized Autoregressive Pretraining for Language Understanding. *Advances in Neural Information Processing Systems*, 32, 5753–5763.
- Yiridoe, E. K., Bonti-Ankomah, S., & Martin, R. C. (2005). Comparison of consumer perceptions and preference toward organic versus conventionally produced foods: A review and update of the literature. *Renewable Agriculture and Food Systems*, 20(04), 193–205. <https://doi.org/10.1079/RAF2005113>
- Yu, C. H., Jannasch-Pennell, A., & DiGangi, S. (2011). Compatibility between Text Mining and Qualitative Research in the Perspectives of Grounded Theory, Content Analysis, and Reliability. *Qualitative Report*, 16(3), 730–744. Retrieved from <https://eric.ed.gov/?id=EJ926322>
- Zablocki, A., Schlegelmilch, B., & Houston, M. J. (2019). How valence, volume and variance of online reviews influence brand attitudes. *AMS Review*, 9(1-2), 61–77.
<https://doi.org/10.1007/s13162-018-0123-1>
- Zenith. (2019). Media Consumption Forecasts 2019. Retrieved from https://www.statista.com/statistics/319732/daily-time-spent-online-device/*, (accessed October 19, 2020).
- Zepeda, L., & Deal, D. (2009). Organic and local food consumer behaviour: Alphabet Theory. *International Journal of Consumer Studies*, 33(6), 697–705.
<https://doi.org/10.1111/j.1470-6431.2009.00814.x>
- Ziegele, M. (2016). *Nutzerkommentare als Anschlusskommunikation: Theorie und qualitative Analyse des Diskussionswerts von Online-Nachrichten*. Wiesbaden: Springer VS.

- Ziegele, M., Breiner, T., & Qiring, O. (2014). What Creates Interactivity in Online News Discussions? An Exploratory Analysis of Discussion Factors in User Comments on News Items. *Journal of Communication*, 64(6), 1111–1138. <https://doi.org/10.1111/jcom.12123>
- Ziegele, M., Springer, N., Jost, P., & Wright, S. (2017). Online user comments across news and other content formats: Multidisciplinary perspectives, new directions. *SCM Studies in Communication and Media*, 6(4), 315–332. <https://doi.org/10.5771/2192-4007-2017-4-315>
- ZMG. (2020). Print and online reach of newspapers in Germany in 2020. Retrieved from <https://www.statista.com/statistics/1114956/print-online-reach-newspapers-germany/>*, (accessed October 19, 2020).

Appendices

Appendix 1 to Section 3.3 Synchronous Topic Correlations.

Let $t = 1, \dots, 10$ be the topic index, T the total number of topics, $i = 1, \dots, N$ the article index, N the total number of articles, $x_{i,t}$ the proportion of topic t in article i , and $y_{i,t}$ the proportion of topic t in all comments of article i . Then, for one topic of interest t , $\rho_t = \rho_{xy_t}$ is the Pearson correlation coefficient of the topic proportions $(x_{i,t}, y_{i,t})$ with $i = 1, \dots, N$.

Appendix 2 to Section 3.3 Cross-Lagged Topic Correlations.

The time dimension was included into the calculation of the Pearson correlation as follows: The years 2007 to 2020 were divided into time windows of two weeks. Within a given two-week time window $w = 1, \dots, W$ with W being the total number of time windows, the number of articles is denoted as x'_w and the number of comments as y'_w . The proportion of topic t in the articles within w is $x_{w,t} = \frac{x'_w t}{x'_w}$, and the proportion of topic t in the comments w is $y_{w,t} = \frac{y'_w t}{y'_w}$. Now, let w_k be an arbitrary two-week time window and w_{k+1} the subsequent time window. Then, we calculated the autocorrelation for articles $P_{X_1 X_2} = \rho_{X_1, X_2}$ based on all value pairs $(x_{w_k, t}, x_{w_{k+1}, t})$ with $k = 1, \dots, W - 1$ and $t = 1, \dots, T$. This was done analogously for the autocorrelation of comments $P_{Y_1 Y_2}$, as well as for the synchronous correlations $P_{X_1 Y_1}$ and $P_{X_2 Y_2}$, where the time-window is always the same.

Appendix 3 to Section 3.3 Normalized Shannon Entropy.

We represented the agendas as time-evolving distributions of topics. We selected a time window length of six months with a step size of two months between consecutive windows. For each time window, we measured the agenda diversity with the normalized Shannon entropy H over the topic distribution (Boydston et al., 2014) (see Eq. 3.1). H measures the diffusion of the distribution by quantifying the level of information. H increases when the topics are equally distributed (i.e., all topics are equally probable, agenda diversity is high), and H decreases when certain topics dominate (i.e., certain topics have high probability, agenda diversity is low). Thus, we can interpret time periods

of low agenda diversity by identifying local minima of H . Building on Tukey's box plot construction, the lower inner fence of H is defined by $LIF = Q_1 - 1.5 \cdot (Q_3 - Q_1)$, with Q_1 as the lower quartile and Q_3 as the upper quartile. Values of H below the LIF indicate local minima, where agenda diversity is low as certain topics dominate. Sensitivity analyses revealed that time windows larger than 6 months were not sensitive enough to show true local minima, and time windows shorter than 6 months delivered artificial minima due to lack of data. Thus, a window length of 6 months was most appropriate. (The same applies to the metric of JSD for agenda distance).

$$H[p] = \frac{-\sum_{i=1}^T (p(x_i) \cdot \ln p(x_i))}{\ln T} \quad \text{Eq. 3.1}$$

Appendix 4 to Section 3.3 Agenda Distance - Jensen-Shannon Distance (JSD).

JSD measures the similarity between two distributions. It is a symmetric and smoothed version of the Kullback-Leibler divergence (see Eq. 4.1). JSD was used to quantify the similarity between topic distributions of all articles (P) and all comments (Q) from a given time period (see Eq. 4.2). The time period specification is the same as for normalized Shannon entropy. We computed the sum of the normalized distributions of all articles in a given time window and re-normalize it. This process was repeated for the comments, and the JSD of the two resulting distributions is calculated. If JSD is close to 0, the topic distributions of articles and comments in this time period are very similar. At large values of JSD, the distributions of articles and comments diverge. Such local maxima of JSD were detected for values of JSD above the upper inner fence (UIF) of the box plot construction, which is defined by $UIF = Q_3 + 1.5 \cdot (Q_3 - Q_1)$, analogously to normalized Shannon entropy.

$$KL(P \parallel Q) = \sum P(x) \left(\log \frac{P(x)}{Q(x)} \right) \quad \text{Eq. 4.1}$$

$$JSD(P \parallel Q) = \sqrt{\frac{1}{2} KL\left(P \parallel \frac{P+Q}{2}\right) + \frac{1}{2} KL\left(Q \parallel \frac{P+Q}{2}\right)} \quad \text{Eq. 4.2}$$

Appendix 5 to Section 3.4 Sample characteristics.

Variables	Sample
N	1118
Gender (%)	
Female/male	50.9/49.1
Age (%)	
18-20	4.7
21-29	14.0
30-49	36.0
50-59	22.5
60-64	9.5
65 and more	13.3
Schooling (%)	
Less than high school	1.8
High school (9 years)	16.5
High school (10 years)	45.9
High school (12/13 years, Abitur)	35.9
Professional education (%)	
No professional education	17.1
Apprenticeship	63.2
Bachelor's degree	4.7
Master's degree/diploma or higher	14.9
Household size (mean)	2.35
Children (< 18 years) in household (mean)	1.45
Household income (%)	
Under 900€	7.6
900–1299€	10.1
1300–1499€	6.4
1500–1999€	10.6
2000–2599€	16.2
2600–3199€	11.8
3200–4499€	16.6
4500–5999€	8.5
6000€ and more	3.8
No answer	8.4

Appendix 6 to Section 3.4 Means (M) and standard deviations (SD) for all self-reported variables (N = 1118).

Variables	Type	M (SD)
Policy acceptance		
“If there was a vote on the following possible political measure, would you vote in favor or against this measure? Receiving agricultural subsidies is to more strongly depend on the compliance with environmental, climate protection, and animal welfare criteria. Thus, in particular organic farmers will receive more support as organic agriculture abides by higher environmental and animal welfare standards compared to conventional agriculture.”	1 = would definitely vote against – 5 = would definitely vote in favor	4.28 (0.93)
Attitude valence		
“How do you feel about organic food?”	-8 = extremely oppose – 8 = extremely favor	3.21 (3.93)
Attitude strength		
	7 items, Cronbach’s alpha = 0.802	27.18 (6.32)
“How important is organic food to you?”	1 = not important – 7 = important	4.64 (1.57)
“How strongly do you identify with organic food?”	1 = not at all – 7 = strongly	4.15 (1.62)
“How representative of your values is your attitude toward organic food?”	1 = not representative – 7 = representative	4.46 (1.63)
“How sure are you of your positions regarding organic food?”	1 = not sure – 7 = sure	4.99 (1.34)
“How informed do you feel about organic food and its production?”	1 = not informed – 7 = informed	4.36 (1.35)
“How likely are you to change your attitude toward organic food?”	1 = unlikely – 7 = likely	4.57 (1.39)
Buying experience		
“For how long have you been buying organic food?”	1 = never – 7 = for more than 5 years	5.08 (1.86)
Media use		
	9 items, Cronbach’s alpha = 0.719	35.73 (9.01)
“How often do you watch TV?”	1 = never – 7 = more than three times a day	5.54 (1.40)
“How often do you listen to the radio?”	see above	4.98 (1.79)
“How often do you read news in print?”	see above	3.87 (1.86)
“How often do you read news online?”	see above	5.12 (1.65)
“How often do you read reader comments on news websites?”	see above	3.65 (1.94)
“How often do you use social networks (e.g., Facebook)?”	see above	4.82 (2.20)
“How often do you create posts in social networks?”	see above	2.80 (1.93)
“How often do you read online question-and-answer forums?”	see above	2.85 (1.74)
“How often do you create posts in online forums or news websites?”	see above	2.10 (1.62)
Media trust		
“If you think of a news media outlet that you use frequently, do you think it can generally be trusted?”	1 = cannot be trusted – 5 = can be trusted	3.50 (1.00)
Theme importance		
“How important is animal welfare with respect to organic food and its production to you?”	1 = not important – 7 = important	5.83 (1.35)
“How important is biodiversity with respect to organic food and its production to you?”	see above	5.20 (1.46)

Appendix 7 to Section 3.4 Priming items.

“Included here are headlines regarding topics that have been publicly discussed recently in Germany. The headlines were picked randomly from news websites by a computer. For each topic, please indicate whether you have read or heard about it.” (answer options: Yes, No, I am not sure)

Weak animal welfare prime:

1. Factory farming in Germany: Only slow progress in animal welfare
2. Why the refugee influx is increasing again
3. Companies are becoming more family-friendly
4. Hard Brexit puts British economy at risk
5. Autonomous driving: Who is liable for traffic accidents?
6. Education experts advocate a centralized school-leaving certificate in Germany
7. Smartphones frequently affected by data security loopholes

Strong animal welfare prime:

1. Factory farming in Germany: Only slow progress in animal welfare
2. Why the refugee influx is increasing again
3. Companies are becoming more family-friendly
4. Animal husbandry: Many animals – little space
5. Autonomous driving: Who is liable for traffic accidents?
6. Education experts advocate a centralized school-leaving certificate in Germany
7. Animal-friendly husbandry: Stricter animal-welfare standards in organic farming

Weak biodiversity prime:

1. Biodiversity: Almost a third of wild plants are endangered
2. Why the refugee influx is increasing again
3. Companies are becoming more family-friendly
4. Hard Brexit puts British economy at risk
5. Autonomous driving: Who is liable for traffic accidents?
6. Education experts advocate a centralized school-leaving certificate in Germany
7. Smartphones frequently affected by data security loopholes

Strong biodiversity prime:

1. Biodiversity: Almost a third of wild plants are endangered
2. Why the refugee influx is increasing again
3. Companies are becoming more family-friendly
4. Organic farming: a first step to more biodiversity
5. Autonomous driving: Who is liable for traffic accidents?
6. Education experts advocate a centralized school-leaving certificate in Germany
7. Why insects are steadily disappearing

Control condition:

1. The number of single households in Germany is increasing
2. Why refugee influx is increasing again
3. Companies are becoming more family-friendly
4. Hard Brexit puts British economy at risk
5. Education experts advocate a centralized school-leaving certificate in Germany
6. Smartphones frequently affected by data security loopholes

De-briefing of participants: The headlines presented were actually published on online news websites. However, they were not randomly selected by a computer. Each participant was randomly attributed to a group of pre-selected headlines.

