Perceived eco-friendliness and perceived quality of consumer products: A theoretical and empirical investigation

Inaugural dissertation

at the Chair of Sales and Marketing

at the Otto-Friedrich-University of Bamberg

submitted by

Katharina Schaper

Bamberg 2026

This work was submitted as a dissertation to the Faculty of Social Sciences, Economics and Business Administration at Otto Friedrich University in Bamberg.

First reviewer: Prof. Dr. Björn Ivens

Second reviewer: Prof. Dr. Susanne Wigger-Spintig

Additional reviewer: Prof. Dr. Eric Sucky

Date of oral exam: 12.01.2026

URN: urn:nbn:de:bvb:473-irb-112536x

DOI: https://doi.org/10.20378/irb-112536

ACKNOWLEDGEMENTS

In this preface, I would like to express my appreciation and gratitude to the most important people along my journey of developing this thesis. I am deeply thankful to all those who supported me along the way.

First and foremost, I would like to express my sincere gratitude to my supervisors, Prof. Dr. Björn Ivens and Prof. Dr. Susanne Wigger-Spintig. Your expertise and guidance have been invaluable throughout this journey. Thank you, for challenging me to reach higher and to grow as a researcher and person and for sticking with me for all this time. I would also like to extend my thanks to Prof. Dr. Eric Sucky as my third reviewer.

I would also like to express my appreciation to my family, friends and colleagues who supported me during the different stages of this project. There are too many to name every one, so I will only mention a few by name, but thinking of many more incredible people that were part of this journey. Special thanks…

- to my partner in crime and significant other, Dr. Patrick Haslanger (for many reasons but most importantly for dealing with me during all ups and downs, for not letting me give up, for continuously challenging me and sometimes also for distracting me)
- to my involved friends along this journey (especially Stefanie Götze, Nadine Wild, Lara Schnappinger)
- to my fellow (formerly) doctoral students (especially Dr. Christopher Neumann, Dr. Lukas Brenner, Dr. Andreas Wild and the Strategy& PhD Crew)
- to my formal and informal mentors (especially Reinhard Vocke, Laura Leeb, Harald Dutzler, Dr. Hans-Jörg Kutschera, Dr. Stephan Gross, Prof. Dr. Philipp Wunderlich)

My deepest gratitude belongs to my family, especially my parents Ursula and Dirk Schaper (as well as my grandparents), who always support me unconditionally. I cannot thank you enough for all you have done and continue to do for me. Your unconditional love, encouragement, and sacrifices have been the foundation of everything I have achieved.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AIC	Akaike information criterion
Ani	Perceived animal friendliness
AVE	Average variance extracted
BIC	Bayesian information criterion
BP	Breusch-Pagan test (statistic)
CO2	Carbon dioxide
CB SEM	Covariance-based structural equation model(ling)
CFA	Confirmatory factor analysis
CFI	Comparative fit index
CI	Confidence interval
Cli	Perceived climate friendliness
CP	Conventional product
CSR	Corporate social responsibility
df	Degrees of freedom
Dur	Perceived durability
DW	Durbin-Watson test
e.g.,	Exempli gratia, for example
EF	Eco-friendliness / eco-friendly
EFA	Exploratory factor analysis
Effort	Perceived producer eco-effort
EFP	Eco-friendly product
et al.	Et alii/ et aliae: And others
etc.	Et cetera: And so forth
EU	European Union
Exper	Perceived producer eco-expertise
f^2	Effect size (Cohens f^2)
FLC	Fornell-Larcker criterion
fsQCA	Fuzzy-set qualitative comparative analysis
GT	Grounded theory
HTMT	Heterotrait-monotrait ratio of correlations
i.e.,	Id est: That is
IAT	Implicit association test
Indic	Shared indicator beliefs
InfObs	Influential observations
Innov	Perceived product innovativeness
KMO	Kaiser–Meyer–Olkin
MAE	Mean absolute error
MAXQDA	Software for qualitative data analysis
MI	Modification index / indices
ML	Maximum likelihood
MLR	Multiple linear regression

(Amazon) MTurk	Crowdsourcing marketplace
N	Sample size
n.s.	Not significant
Nat	Perceived naturalness
p	Probability of error
p.	Page
PEF	Perceived eco-friendliness
PEF_dim	Multidimensional measure of PEF
PEF_uni	Unidimensional measure of PEF
PEFP	Perceived eco-friendly product
PI	Purchase intentions
PLS SEM	Partial least square structural equation model(ling)
pp.	Pages
PQ	Perceived quality
PQR	Perceived quality risk
PV	Perceived value
Q-Q plot	Quantile-quantile plot
r	Correlation coefficient
R	Software environment for statistical computing and graphics
R^2	Explained variance
Rec	Perceived reuse- and recyclability
Rep	Perceived repairability
Res	Perceived resource efficiency
RMSE	Root-mean-square error
RMSEA	Root mean squared error of approximation
RRP	Remanufactured / refurbished product
RSB	Resource synergy beliefs
SD	Standard deviation
SE	Standard error
SEM	Structural equation model(ling)
SMSR	Standardized root mean square residual
Soc	Perceived social friendliness
Stren	Strength beliefs
TLI	Tucker Lewis index
UN	United Nations
USA	United States of America
VHB	Association of German professors in business research (“Verband der Hochschullehrer für Betriebswirtschaft e.V.”)
VIF	Variance inflation factor
vs.	Versus
W	Shapiro-Wilk test (statistic)
Wglow	Anticipated warm glow
WTP	Willingness to pay

1. Introduction

1.1. Motivation and background

With the effects of climate change and global warming substantially impacting human life and wildlife (European Commission, 2016), eco-friendly products as means to reduce or reverse the negative effects of consumption on the environment are receiving increasing attention from consumers, firms and public bodies (Barbarossa & De Pelsmacker, 2016; Bhardwaj et al., 2020; Owen et al., 2018; W. Wang et al., 2017). For the purpose of this thesis, an objectively eco-friendly product is defined as a product that minimizes its direct and indirect impact on the environment (including ecological and social aspects) during its whole lifecycle, compared to earlier versions of the same product or similar products that fulfill the same consumer need.

In line with the attention towards those products, consumption is a relevant lever to tackle environmental and social problems, because issues like resource depletion, chemical pollution, toxic waste, biodiversity loss, lack of freshwater or poverty are caused or accelerated by human consumption (Rost, 2015). Political institutions like the *United Nations* [UN] (with the Sustainable Development Goals), climate activists like Greta Thunberg and repeated reports of natural disasters like wildfires and flooding brought global attention through mass media coverage to this topic. Consequently, consumer awareness of climate change has increased (Dutta-Powell et al., 2024). In line with that, most Europeans believe it is very important for them personally to protect the environment and two-thirds say that their consumption negatively affects the environment (European Commission, 2020). Research has shown that the broad coverage of environmental issues in media can help fueling sales of *perceived eco-friendly products* [PEFPs] (Y. Chen et al., 2019). However, overall sales of these products still remain low: While 65% of consumers say they plan to purchase products from sustainable brands, only 26% actually purchase them (K. White, Hardisty, et al., 2019). Similarly, according to the key indicator of the German Environmental Agency, market shares of products with official sustainability labels are increasing but were only at 7.6% in 2015 whereas the goal is to reach 34% until 2030 in Germany (Steinemann et al., 2017). In 2022, the market share in Germany was 12.2% for products with official sustainability labels, with big differences between product types: While organic food had a market share of 6.3%, washing machines with highest energy efficiency had a share of 95.6%, yet electric cookers with the same energy efficiency class had < 1% market share (German Environmental Agency, 2024).

Overall, even though consumers have a positive attitude towards PEFPs this does not necessarily translate into action, i.e., actual PEFP purchase or other ethical behaviors like recycling (ElHaffar et al., 2020). This well-documented phenomenon is called "attitude behavior gap". Research has identified a number of barriers creating this gap, including prices, quality, perceived risks, social and physical context, lacking knowledge, lacking trust, lacking self-efficacy, altruism, locus of control and many more (Derdowski et al., 2020; ElHaffar et al., 2020; A. P. Sharma, 2021). Important aspects related for example to consumer knowledge and trust, are challenges for consumers to identify the respective products: For example, consumers are often not sure whether to believe in the *eco-friendliness* [EF] of a given product as claimed by marketing efforts because greenwashing is a common concern (Martínez et al., 2020). Greenwashing means spreading incomplete or false information about the environmental performance of a product or a firm as a whole to create an eco-friendly image, thus misleading consumers (Becker-Olsen & Potucek, 2013; De Freitas Netto et al., 2020; Delmas & Burbano, 2011; Furlow, 2010). More and more companies and products advertise themselves as being *eco-friendly* [EF] (Lyon & Montgomery, 2015), while this is by far not always true (Kahraman & Kazançoğlu, 2019; Martínez et al., 2020). With consumers being increasingly aware of greenwashing (Rahman & Nguyen-Viet, 2023), it is of interest to understand what they currently perceive as eco-friendly.

In addition, research has shown that consumer perceptions and "actual" eco-friendliness as measured via life-cycle-analyses can differ substantially (Boesen et al., 2019; Steenis et al., 2017). For this thesis, the consumer perspective is taken, meaning products that consumers perceive to be eco-friendly are considered, regardless of their objective eco-friendliness. Assuming a consumer-centric approach acknowledges these differences and focuses on what drives consumer decisions. Taking this consumer perspective might also help to explain the gap between interest in *eco-friendly products* [EFPs] and actual consumption. Consumers perceive PEFPs as different from *conventional products* [CP] on several dimensions, for example regarding their gentleness / strength: Luchs et al. (2010) explored consumer associations and preferences and found that more sustainable products are associated with a higher degree of gentleness; thus, when strength-related attributes are valued (e.g., in laundry detergents or car tires), consumers can prefer CPs over PEFPs.

Perceived eco-friendliness [PEF] can not only influence gentleness perceptions, but also consumers' overall quality perceptions. *Perceived quality* [PQ] is "the customer's perception of the overall quality or superiority of a product or service with respect to its intended purpose

relative to alternatives" (Aaker, 1991, p. 85). PQ is a highly relevant concept because it does affect other theoretical and behavioral variables like value perceptions, willingness to pay, product choice, buying behavior or product usage (De Pelsmacker & Janssens, 2007; Homburg et al., 2005; Lin & Chang, 2012; Van Doorn & Verhoef, 2011). Positive relationships between PQ and *purchase intentions* [PI] have been found for many products, e.g., food (Zheng et al., 2018), apparel (Tong & Su, 2018), and recycled products (H. Sun et al., 2018). PI describe the willingness and likelihood of consumers towards purchasing a certain product. Overall, the relevance of PQ for PI has been reported for decades within all kinds of contexts (Zeithaml, 1988), making PQ a key concept for research and praxis.

Because PQ is such a central concept, several authors have investigated which effect PEF of a product has on quality perceptions. When looking specifically at the correlations between PEF and PQ, some researchers argue that consumers perceive PEFPs as of higher quality than CPs (e.g., X. Chen et al., 2018; Frank & Brock, 2019; Magnier et al., 2019; Steinhart et al., 2013), but several others state the opposite (e.g., Elgaaïed-Gambier, 2016; Gleim & Lawson, 2014; Kuah & Wang, 2020; Lin & Chang, 2012; Newman et al., 2014). While these positions may seem to be contradictory, various relationships between PEF and PQ can co-exist. The term *relationship* as used in this thesis describes "an association or connection between objects, events, variables, or other phenomena" (American Psychological Association, 2018a). Examples for factors contributing to varying PEF-PQ relationships include: Consumers evaluate different products differently, e.g., high or low priced (Bodur et al., 2016), different consumers evaluate the same products differently, e.g., women or men (Van Doorn & Verhoef, 2011), or PEF can be displayed differently, e.g., via verbal claims, product material, or design.

As previous research and practical examples have shown, PEF and PQ are concepts warranting further attention in research. The next sections will outline the objectives, scope and then structure of this thesis for investigating PEF and PQ.

1.2. Thesis objectives and scope

As outlined in the previous section, there are multiple studies assessing PQ of PEFPs with differing products, differing research methods, and differing resulting findings. However, an overview of the correlations between PEF and PQ is not available to date. While a growing number of literature reviews addressing EF in the field of marketing has been published in the past few years, they address other aspects of sustainable consumption like shifting consumer behavior (Rana & Paul, 2017; K. White, Habib, et al., 2019), theories explaining green

consumer behavior (Groening et al., 2018), green marketing mix (Dangelico & Vocalelli, 2017), circular economy (Bigliardi et al., 2020; Camacho-Otero et al., 2018; Geissdoerfer et al., 2017; Wagner & Heinzel, 2020), or consumer perceptions of eco-friendly packaging (Boz et al., 2020; Ketelsen et al., 2020; Otto et al., 2021). Other recent reviews focus only on specific industries like cosmetics (Bom et al., 2019; Liobikienė & Bernatonienė, 2017), fashion (Wagner & Heinzel, 2020) or wine (Schäufele & Hamm, 2017).

In addition, despite the growing body of research involving PEF (Bhardwaj et al., 2020), little is known about how consumers define and form perceptions of these products (Gershoff & Frels, 2015). Several authors (e.g., Durif et al., 2010; Sdrolia & Zarotiadis, 2019) have provided definitions of green products (the term being often used interchangeably with eco-friendly, as well as sustainable, ethical or environmental-friendly), mainly from a scientific perspective. To the best of the author's knowledge, no one has worked directly together with consumers to holistically understand their perceptions of what PEFPs are.

The key objectives of this thesis are to a) conceptualize PEF as understood by consumers and b) provide an extensive overview of the relationships between PEF of consumer products and the associated PQ. In doing so, it intends to summarize existing research and add new insights via qualitative and quantitative primary research and analyses. Specifically, a three-step approach consisting of a systematic literature review (see chapter 3), an interview-based qualitative exploration (see chapter 4) and an online survey-based quantitative investigation (see chapter 5) is applied to reach the objective. Using multiple methods allows to triangulate findings and thus present stronger results, address broader research questions, create a more comprehensive picture and provide a holistic understanding of the topic under investigation (Davis et al., 2011).

The scope of this thesis is limited to tangible consumer products in a business to consumer setting across industries. In this thesis, the term *industry* describes a group of products with similar areas of application and purpose, e.g., food industry or clothing industry. This thesis does not consider intangible services (like hospitality services) or product perceptions in business-to-business settings. The decision to exclude services was taken due to the differences between tangible products and intangible services. For example, different definitions of PQ of a product or of a service are used in the literature (De Oña & De Oña, 2015; Golder et al., 2012) and consumers may evaluate PQ of products versus services differently (Walsh & Beatty, 2007). Tangible products were chosen because an initial screening showed more research about PEFPs and PQ than about eco-friendly services, emphasizing the importance of PQ for PEFPs

and the need for synthesizing findings. The tangible attributes linked to PEFPs like product colors, labels, ingredients, materials, etc. provide many options for different perceptions and correlations between PEF and PQ, which cannot be directly applied to services.

Based on the objectives as described above, these overarching research questions will be investigated throughout the thesis with each chapter focusing on specific aspects:

1) How are PEFPs defined and conceptualized?

2) What are the relationships between PEF and PQ?

For the qualitative and quantitative studies conducted within this thesis, their individual research questions (derived from these overarching questions) will be explained within the respective chapters as outlined in the next section about thesis structure.

1.3. Thesis structure

This first chapter so far has provided a general introduction to the topic of PEF and its relevance for research and practice. The objectives of this thesis have been explained, and the scope has been specified. All chapters add specific elements to answer the overarching research questions as presented in the previous section.

Chapter 2 lays the theoretical groundwork of this thesis. It describes and defines the key concepts of perception, PEF and PQ. For perception, the key processes involved are outlined. As focal construct of this thesis, different definitions of PEF and an overview of key words like sustainable, green, social, ethical, etc. are presented before the working definition for this work is given. For PQ, main definitions in the literature are summarized and the high-level quality perception process outlined, before the working definition for this thesis is given.

Chapter 3 systematically investigates the status quo of research on relationships between PEF and PQ. The literature review followed the process and suggestions for systematic reviews as described by Palmatier et al. (2018). First, the selection approach for relevant articles is described. Then, descriptive results (publications over time, occurrence in journals, summary of research designs and PQ measures) are shown. The operationalizations of PEF in the reviewed articles are presented before a detailed section about the relationships between PEF and PQ relationships. This section includes results on the direct effects (including a high-level industry summary), covariates (in terms of product, firm, consumer, and situation related variables) and spillover effects between PEF and PQ. In the last parts of this chapter, findings and gaps for future research are discussed and limitations outlined, followed by a brief

conclusion. After chapter 3, the primary data collection and research commences applying mixed methods research.

Chapter 4 describes the qualitative research conducted to develop a theoretical model grounded in data. In the beginning of the chapter, the specific aim and research question for that part of the thesis is described. Thereafter, grounded theory as research method is introduced in detail. Examples of the analyses and coding procedures are shown. Following, results of this research are presented: PEF of consumer products is conceptualized as multidimensional construct. Correlations with PQ and covariates are explored and synthesized, being clustered into perceiver, product, and producer beliefs. The hypotheses for further quantitative investigation are specified. A discussion with theoretical and managerial implications as well as limitations, and a brief summary conclude the chapter.

In chapter 5, the theoretical model as developed in chapter 4 is further assessed quantitatively. After introducing the specific aim and research questions, the quantitative research methods are introduced. Pretests that were conducted to ensure proper and relevant measurement are described before the design and measurement are explained. The final sample and specifics of the data collection are presented. Then, the key methods for data analyses (i.e., factor analysis, regression and structural equation modelling) are introduced. Afterwards, the results of all analyses are presented. The proposed dimensions of PEF are assessed and compared with a unidimensional measure. The direct effect of PEF on PQ is assessed for two different products and the impact of the proposed covariates for each product described. Next, the implications of the results are discussed and limitations outlined before a summary closes this chapter.

Finally, critical reflections of the overall work are shared in chapter 6. Results across the different parts of this thesis are brought into perspective and their implications for research and practice are laid out. Limitations of the thesis as a whole are outlined and connected to future research. At the end, a summary connecting all parts concludes this work.

Figure 1 visualizes the key content elements of this thesis as described above. Sections at a higher level that are not included in the Figure (e.g., 3.4) have an introductory and summarizing function and are thus not visualized here separately.

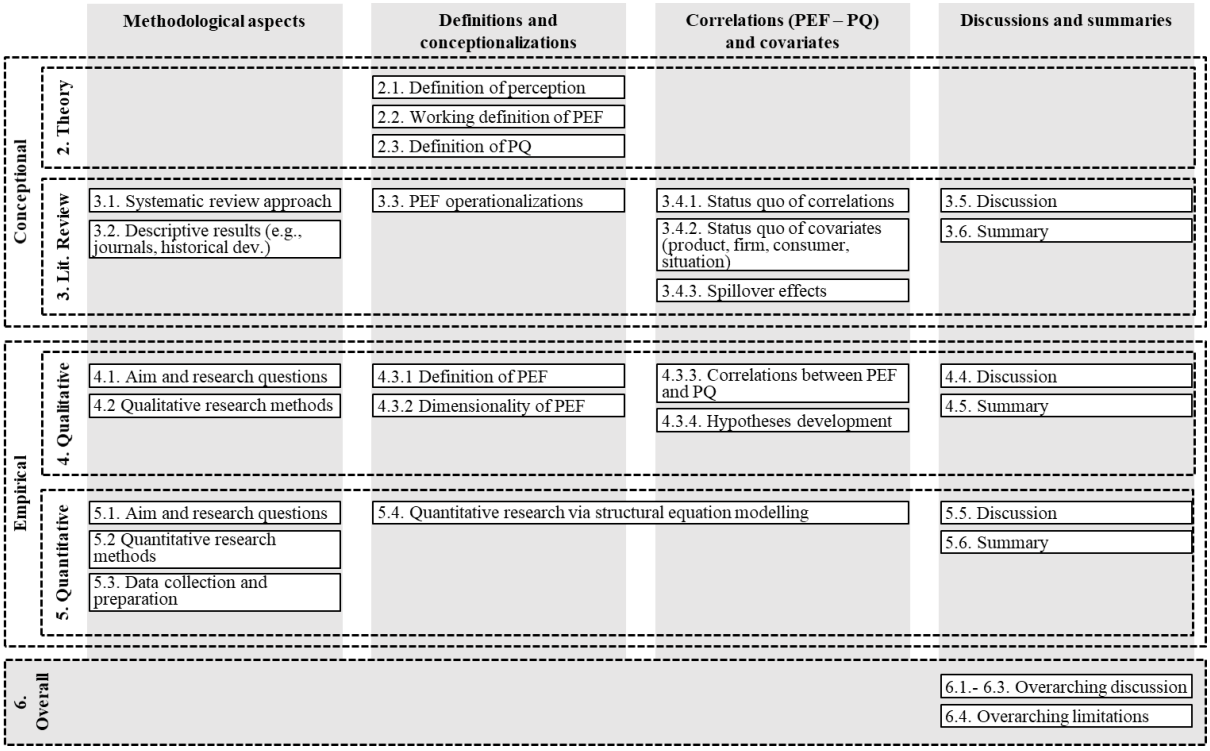


Figure 1: Structure and key components of this thesis.

2. Theoretical basis

In this section, definitions of the key concepts for this thesis are outlined and differences with related concepts are described to determine the scope of the analysis. The concepts of perception, PEF / PEFP and PQ are described and explained. For this thesis, PEF and PQ are considered as separate concepts in line with definitions in the literature (see sections 2.2 and 2.3). Table 1 shows a summary of the working definitions used.

Table 1: Working definitions of key concepts for this thesis.

Term	Definition used
Perception	The process or result "of organizing and interpreting sensory information, enabling us to recognize meaningful objects and events" (Myers & DeWall, 2015, p. 230)
Perceived eco-friendly product (PEFP)	A product that consumers perceive to minimize its direct and indirect impact on the environment (including ecological and social aspects), compared to earlier versions of the same product or similar products that fulfill the same consumer need (based on Sdrolia & Zarotiadis, 2019).
Perceived quality (PQ)	"The customer's perception of the overall quality or superiority of a product or service with respect to its intended purpose relative to alternatives." (Aaker, 1991, p. 85)

2.1. Perception

Perception has been a central concept in fields like philosophy and psychology for the past centuries and millennia. According to authors like Gaukroger (2000), Grene (1974), Hwang (2011) or Slakey (1961), famous philosophers like Aristotle, Socrates or Descartes studied and wrote about perception. From the philosophic stance, two opposing viewpoints led to discussions between "'direct realists', who see perception as immediate awareness of the world, and 'indirect' or 'representative realists', who argue that there can be no such awareness without the involvement of some kind of mediating representation in cognitive processing" (Sutton, 2000, pp. 524–525). Cognitive processing is defined as "all the mental activities associated with thinking, knowing, remembering, and communicating" (Myers & DeWall, 2015, p. 186). Thus, representative realists argue that thoughts, knowledge, memories, etc. play a vital role for perception. In line with modern psychology, the latter view is taken in this thesis.

Previous literature presents multiple definitions of perceptions. According to the respective authors, perception is defined as:

- "the process of organizing and interpreting sensory information, enabling us to recognize meaningful objects and events." (Myers & DeWall, 2015, p. 230)
- "the organization, identification, and interpretation of a sensation in order to form a mental representation." (Schacter et al., 2011, p. 127)
- "the process through which people take raw sensations from the environment and give them meaning, using knowledge, experience, and understanding of the world." (Bernstein & Nash, 2008, p. 86)
- "an individual's or group's unique way of viewing a phenomena; involving the processing of stimuli; and incorporating memories and experiences in the process of understanding." (S. M. McDonald, 2012, p. 8)
- "conscious and unconscious effects of sensory stimulation on behavior." (Norman, 2002, p. 73)

This thesis follows the definition of Myers and DeWall (2015, p. 230) because their work is highly recognized and extensive, is more recent than others and seems well suited for application in the context of PEF and PQ. It extends the definition not only on the process itself but also on the results or outcomes of the perception process. The objects and events that are referred to in the definition of perception do not need to be physical in itself but can also be hypothetical constructs (e.g., fear, intelligence); meaning that they cannot be directly observed (Eid et al., 2011). For example, when someone reports perceiving disgust, it means this person has organized and interpreted the sensory information and recognized to perceive disgust. The same can be said about perceiving EF or quality of a product: After organization and interpretation of sensory information, we recognize and can communicate our perceptions.

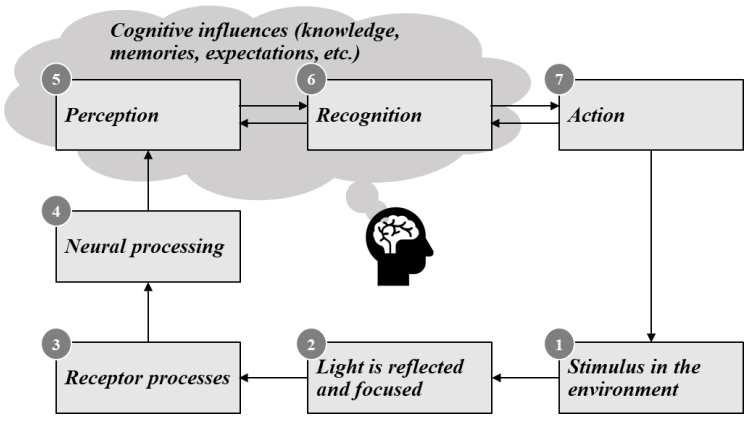


Figure 2: Illustration of the perception process.
Source: Based on Goldstein & Brockmole (2017).

The preception process (see Figure 2) follows seven steps according to Goldstein & Brockmole (2017), from a stimulus in the environment to some kind of action. Even though this is a simplified version, it is helpful to get an overview and understand the concept of perception. The process is pictured as a circle because the action alters the situation and thus makes room to perceive an altered or new stimulus. Action is insofar important in research about perception because the perceiving person needs to act (e.g., via verbal answers, scale ratings, etc.) to let the researcher know about their perceptions. The overall process, can be clustered in two major segments: Bottom-up and top-down processing (Myers & DeWall, 2015). Bottom-up processing includes the initial steps (2-4) from the sensory receptors to the information integration in the brain. Top-down processing includes "the higher-level mental processes, as when we construct perception drawing on our experience and expectations" (Myers & DeWall, 2015, p. 230). This means that cognition, defined as "all the mental activities associated with thinking, knowing, remembering, and communicating" (Myers & DeWall, 2015, p. 186), has an important influence. Top-down processing can happen with or without being aware of it (Goldstein & Brockmole, 2017). The transformation between these two segments (between step 4 and 5) "is perhaps the most miraculous of all, because electrical signals (Step 4) are transformed into conscious experience: The person perceives the tree (Step 5) and recognizes it (Step 6). We can distinguish between perception, which is conscious awareness of the tree, and recognition, which is placing an object in a category, such as 'tree,' that gives it meaning" (Goldstein & Brockmole, 2017, p. 8).

In bottom-up and top-down processing, perception not only depends on a stimulus, but also on the perceiving individual, making it a highly subjective experience (S. M. McDonald, 2012):

- Bottom-up: Perceptions depend on the specific organ of perception (e.g., eyes) which can vary between individuals. For example, persons with specific alterations in the organ, e.g., fewer cone pigments in the eyes (color blindness), perceive colors different compared to persons with all the typical cone pigments (Goldstein & Brockmole, 2017). Another example is that different species (like frogs) receive and process (visual) stimuli different from humans and have thus different perceptions (Myers & DeWall, 2015). Similar examples can be found for other senses (Myers & DeWall, 2015).
- Top-down: Cognitive influences play a key role for perception and recognition. Factors like knowledge, experiences, expectations, motivation and emotions can have an influence (Goldstein & Brockmole, 2017; Myers & DeWall, 2015). These cognitive

influences are highly individual and differ between people. In addition, individual differences in the brain, for example caused by a tumor, can lead to altered or inhibited perception or recognition. For example, Sacks (1985, as cited in Goldstein & Brockmole, 2017) describes a case where a person was unable to recognize objects (e.g., people or a glove) even though he could see and perceive the parts of it. Thus, the categorization of the object parts did not work for this person.

Because perceptions are individual experiences, the distinction between perceptual and physical (objective) needs to be made. Physical measurements and reported perceptions are often different (Goldstein & Brockmole, 2017). This thesis focuses on consumer perceptions. After this introduction to perception, the concepts of PEF and PEFPs are explained next.

2.2. Perceived eco-friendliness (PEF) and perceived eco-friendly products

The term *eco-friendliness* as noun belonging to the adjective *eco-friendly* is used to describe that something is eco-friendly, in this case consumer products. Thus, eco-friendliness is seen in this thesis as product attribute while eco-friendly behaviors that are not associated with the product itself are not considered (e.g., planting a tree or donating a certain amount of profit). The degree of PEF makes a consumer product being perceived as a more or a less eco-friendly product. The term eco-friendly is often used interchangeably with environmentally friendly, ecological, green, sustainable or ethical in research (e.g., De Medeiros et al., 2014; Gleim et al., 2013; Gleim & Lawson, 2014; Lin & Chang, 2012; Liobikienė & Bernatonienė, 2017; Sdrolia & Zarotiadis, 2019; W. Wang et al., 2017), which all have been considered in this thesis. Consumers also use several terms and struggle to distinguish between concepts like fair trade, local trade, eco-products or organic products (Davies & Gutsche, 2016).

According to Boks and Stevels (2007) eco-friendliness can be analyzed from at least three perspectives:

- Scientific, meaning that the environmental impact is determined as objectively as possible, often using lifecycle assessment (which still bears the risk of not properly considering all relevant aspects),
- Government, meaning priorities based on various factors like geographical position, availability of resources, economic status, and others, also including issues that are important to the voting public (which might not be in line with the scientific view) and

- Consumer, meaning the perceptions that the public has, which are often linked to emotions and can be related to recent events (e.g., after material shortages, more attention is brought to resources.

The consumer perspective is taken for this thesis to consequently focus on the consumer as primary research object. Interest in consumer perceptions of EF has been shown in academic literature: Several authors (Henninger et al., 2016; Magnier & Crié, 2015; Mukendi et al., 2020; A. T. Nguyen et al., 2020; Van Bussel et al., 2022) addressed the topic of defining PEF from consumer perspective within specific industries / product groups. For example, van Bussel et al. (2022) investigated food-related consumer perceptions of EF. Using GT, they identified six domains (i.e., production, transportation, product, product group, consumer, waste, and contextual factors), with corresponding categories and subcategories that consumers consider. However, they state consumers cannot easily define sustainability and do not include a definition in their article. Similarly, Henninger et al. (2016) state that there is no clear definition for sustainable fashion and propose a matrix that producers can leverage to communicate about PEF to consumers to influence their perceptions. Via a GT approach, they find that PEF of fashion products depends on the person and on the context. Mukendi et al. (2020, p. 2877) used a definition of perceived sustainable fashion as "the variety of means by which a fashion item or behavior could be perceived to be more sustainable, including (but not limited to) environmental, social, slow fashion, reuse, recycling, cruelty-free and anti-consumption and production practices". Two other research articles (Magnier & Crié, 2015; A. T. Nguyen et al., 2020) focused on sustainable packaging and started with the definition by the Sustainable Packing Coalition. A.T. Nguyen et al. (2020) then identified three key dimensions of PEF packaging from consumers perspective: Material, manufacturing technology and market appeal. They also defined several dimensions each, e.g., being biodegradable or non-toxic for packaging material. Magnier and Crié (2015) created a taxonomy of PEF packaging cues and also identified three key dimensions: Structural, graphical/iconic and informational cues. While all these studies provide valuable insights, they do not answer what consumers perceive holistically as PEFPs as they are focusing on either specific product types (e.g., fashion products) or on specific product attributes (e.g., packaging).

Overall, looking at the definition of (perceived) eco-friendly products within academic literature, not one single, widely used definition prevails but rather many definitions focusing on different aspects are used. A rather comprehensive overview of definitions can be found in the reviews by Sdrolia and Zarotiadis (2019), and Durif et al. (2010). Durif et al. (2010) looked

at the concept of EF from three perspectives: First, they reviewed 35 academic definitions from 1971 to 2009; Second, they included an industrial perspective through a bibliographic approach; Third, they adopted the consumer perspective (perceptions) via a consumer survey. The consumer survey generated four PEF factors: "(1) non-toxic for nature; (2) good for health; (3) socially responsible; and (4) good for the planet" (Durif et al., 2010, p. 31). However, the study focused on cleaning products only and used predefined statements rather than open questions. In another literature review, Sdrolia & Zarotiadis (2019) focused on the academic perspective with a broader and more recent literature search. They found 51 definitions of EFPs from 1975 to 2017 and described the need for and provided a holistic definition: "Green is a product (tangible or intangible) that minimizes its environmental impact (direct and indirect) during its whole life-cycle, subject to the present technological and scientific status" (Sdrolia & Zarotiadis, 2019, p. 164). Yet this definition does not account for consumer perceptions, and it does not clarify if social impact is considered (i.e., included in environmental impact) or not.

Table 2 gives an overview of selected definitions that were chosen because: 1) Either they are the outcome of a literature review; 2) Have been cited frequently; 3) Or were included in one of the articles identified via the structured literature search (see section 3.1 for details on the search). A working definition for the current work that addresses these points is used as presented later in this section.

Table 2: Selected definitions of EF and related terms in the literature.

Term	Definition	Source
Sustain-able	"Sustainable – to ensure that it meets the needs of the present without compromising the ability of future generations to meet their own needs". "Sustainable global development requires that those who are more affluent adopt life-styles within the planet's ecological means". "Sustainable development is not a fixed state of harmony but rather a process of change in which the exploitation of resources, the direction of investments, the orientation of technological development, and institutional change are made consistent with future as well as present needs".	(UN Secretary-General & World Commission on Environment and Development, 1987, p. 25)
	"Sustainability is the interlinking pillars of the triple bottom line: economics, environment, and society (otherwise known as people, planet, and profits)".	(Elkington, 1994 as cited by Lunde, 2018, p. 89)

Term	Definition	Source
	"Sustainability has two underlying principles. First, only using the earth's resources at a rate which allows them to be regenerated, or (in the case of non-renewable resources) which allows sustainable substitutes to be developed. Secondly it involves creating waste at a rate that can be assimilated by the environment, without impairing it".	(Peattie, 1999, p. 133)
	"Products that have a positive social and/or environmental impact".	(Luchs et al., 2010, p. 18)
	"Products that have a positive environmental and/or social impact because they are produced with concern for human and natural resources, such as air, water, and land".	(Brach et al., 2018, p. 254)
	"A sustainability focus suggests that utilized resources can be renewed by mimicking the circular flows of resources in nature, and it respects the fact that capacity of both resources and the environment are limited". Sustainable consumption results "in decreases in adverse environmental impacts as well as decreased utilization of natural resources across the lifecycle (…)".	(K. White, Habib, et al., 2019, pp. 23–24)
Ethical	Ethical attributes "are related to a variety of social issues (e.g., fair labor practices, humane treatment of animals) and environmental issues (e.g., recycling, avoiding pollution). Although prior literature has often referred to 'ethical attributes' and 'ethical products,' the term 'sustainable' is commonly used in industry practice".	(Luchs et al., 2010, p. 19)
	Products "that address environmental and human-welfare issues".	(Hindsley et al., 2020, p. 1)
Green	A product whose "environmental and societal performance, in production, use and disposal, is significantly improved and improving in comparison to conventional or competitive products offerings".	(Peattie, 1995 as cited by Dangelico & Pontrandolfo, 2010, p. 1609)
	Products that "use less resources, have lower impacts and risks to the environment and prevent waste generation already at the conception stage".	(Commission of the European Communities, 2001, p. 3)
	"A product whose design and/or attributes (and/or production and/or strategy) use recycling (renewable/toxic-free/biodegradables) resources and	(Durif et al., 2010, p. 27)

Term	Definition	Source
	which improves environmental impact or reduces environmental toxic damage throughout its entire life cycle".	
	“A green product is one that is produced with concern for the physical environment: air, water, and land”.	(Gleim et al., 2013, p. 45)
	“A product (tangible or intangible) that minimizes its environmental impact (direct and indirect) during its whole life-cycle, subject to the present technological and scientific status”.	(Sdrolia & Zarotiadis, 2019, p. 164)
	“Products that are intentionally designed to decrease environmental impact with respect to the status quo”.	(Maccioni et al., 2019, p. 1)
	“Products that seek to protect or enhance the environment during production, use, or disposal by conserving resources and minimizing the use of toxic agents, pollution, and waste”.	(Bhardwaj et al., 2020, p. 1)
	“A green product can be defined as a product that has been produced, delivered, and consumed with concerns about the negative externalities that such activities can have with regards to the environment, animals, and society”.	(Dixon & Mikolon, 2021, p. 1)
Environ-mental / environ-mentally friendly	Products “that provide greater environmental benefits, or that impose smaller environmental costs, than similar products”. “Alternatively or in addition”, these products are produced “in ways that are less environmentally burdensome than the production processes of its competitors”.	(Reinhardt, 1998, p. 46)

Key similarities between these definitions are that the whole lifecycle from product design to waste management is considered and that the focus is on the environmental impact (even though social and sometimes economic aspects are also mentioned). Several definitions agree that both the decreased adverse impact as well as the improved benefits are underlying mechanisms for PEFPs (Bhardwaj et al., 2020; Durif et al., 2010; Reinhardt, 1998).

Apart from the similarities between definitions, there are several challenges associated with the definitions of PEFPs in the literature. First, most definitions do not specify what constitutes the reference or benchmark of being considered eco-friendly versus conventional. In the selected definitions shown in Table 2, only the definitions provided by Reinhardt (1998)

and Peattie (1995 as cited by Dangelico & Pontrandolfo, 2010) give a reference: Similar (conventional) products or products / processes of competitors. Even though Maccioni et al. (2019) refer to the 'status quo' it is not clear what that really means. Related to this benchmarking challenge, in most of the definitions it seems that once a product is considered a PEFP, it will remain that way. However, it is rather subject to a process of change (UN Secretary-General & World Commission on Environment and Development, 1987). Addressing this issue, Sdrolia and Zarotiadis (2019) include an important remark regarding the present status. This implies that products that are considered eco-friendly today might not be considered eco-friendly in the future, for example when new technologies allow for further environmental impact reduction. Second, not all articles use the terms (sustainable, green, environmentally friendly, eco-friendly, etc.) in the same way. While many authors use them interchangeably and switch between several terms in their articles, some disagreement of the scope remains. For example, Lunde (2018, p. 86) states in his systematic review of sustainability in marketing that "there is no clear definition of sustainability in marketing. In marketing literature, some researchers define sustainability as an environmental concept, while others define it more holistically".

To clarify the relations between the various key words used in the literature, they have been analyzed and visualized in Figure 3. The visualization has been developed based on the use of key words in the definitions shown above and the application of key words in the studies that resulted from the structured literature search process. Many authors (e.g., Achabou & Dekhili, 2013; Anagnostou et al., 2015; Gassler et al., 2016; Herédia-Colaço & Coelho do Vale, 2018; Iyer & Reczek, 2017; Mai et al., 2019; Mancini et al., 2019; Nichols et al., 2019; Ozanne et al., 2016; Ross & Milne, 2020; H. Wang et al., 2019) agree that 'sustainable' includes both the environmental / ecological and the social aspect as well as an economic aspect (the latter not being relevant for this thesis). Similarly, *ethical* is spanning both ecological and social aspects, yet without the additional economic perspective (e.g., Bodur et al., 2016; Das et al., 2020; Frank & Brock, 2019; Herédia-Colaço & Coelho do Vale, 2018; Hindsley et al., 2020; Luchs et al., 2010). In most studies *green* focuses strongly on the natural environment (e.g., Frank & Brock, 2019; Iyer & Reczek, 2017; Lin & Chang, 2012; Royne et al., 2012; Suki, 2016; Van de Velde et al., 2009) while few emphasize both natural environment and social aspects (e.g., Dixon & Mikolon, 2021; Newman et al., 2014; Tong & Su, 2018). *Eco-friendly* is chosen as the leading term for this thesis because it is rarely defined in literature, allowing to develop and present the most suitable definition for this article without being tied to

preconceptions or disagreements of the term. In addition, consumers are familiar with the term *eco-friendly*, according to Campbell et al. (2015) even more than with the term *sustainable*. Not only from consumer perspective, but also in (high-quality) literature, the term *eco-friendly* is commonly used (e.g., Blasi et al., 2020; Brough et al., 2016; Moons et al., 2018; A. T. Nguyen et al., 2020; Pahlevi & Suhartanto, 2020; Perez et al., 2020; Sundar & Kellaris, 2017); just like the term *eco-label* (e.g., Darnall et al., 2018; Dekhili & Achabou, 2015; Gorton et al., 2021; Hille et al., 2018; Lim & Reed, 2020; Preziosi et al., 2019; Testa et al., 2015), thus showing the relevance of the PEF concept.

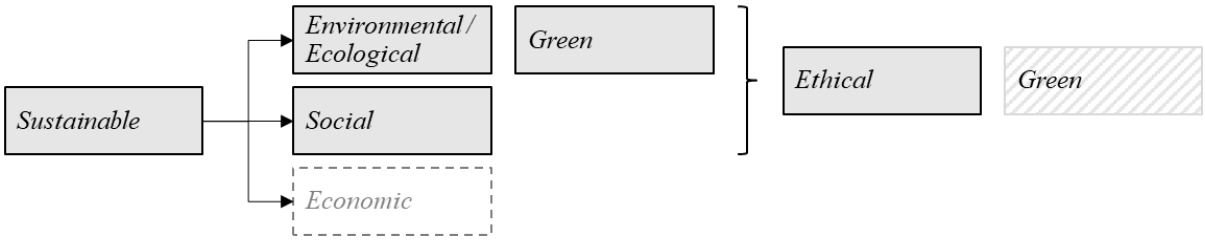


Figure 3: Visualization of key word relationships as mainly used in the academic literature. Notes: The economic perspective of sustainability is not emphasized here (thus appears faded). The term "green" is mainly used to reflect environmental aspects but can occasionally include social aspects (thus included twice and slightly faded in the less common application).

Figure 3 makes the differentiation of product benefits for natural and social environment in the definitions evident. However, this split is not always relevant. For example, fair trade products are perceived to use less chemicals or have more rigid standards (Davies & Gutsche, 2016) which is a benefit for nature and humans involved. Apart from advantages for the natural environment, EF can lead to improved social impact and / or well-being of people (Peattie, 1999; K. White, Habib, et al., 2019). For example, by using less toxins during production, the well-being and health of factory workers can be positively impacted as well as the living conditions and food sources for people living nearby production facilities. As another example, by using biodegradable materials instead of plastic, less plastic waste and micro plastic is set free which would negatively impact animals but also humans (Prata et al., 2020; Smith et al., 2018). There are many other examples - like limiting emissions to slow climate change which would negatively affect humans via health impacts (Wu et al., 2016) or food security and hunger (Chakraborty & Newton, 2011).

Many articles that have defined EFPs took an objective perspective (rather than taking the consumer view). This thesis first also provides a definition of objectively defined eco-friendly products and then develops a definition of PEFPs. In doing this, the definition by Sdrolia and Zarotiadis (2019) as outcome of their systematic literature review is used as basis

because it is one of the most recent and thoroughly developed definitions. Two aspects are added in the definition used here to address the identified challenges of existing definitions: The specification of ecological and social impacts and the clarification what is used as a reference for comparison. Clarifying these aspects is important to be able to precisely apply the definition. Additionally, in line with the scope of this article (as described in the introduction), only tangible products are considered. Thus, the developed definition of objectively EFPs used in this thesis is the following:

An objectively EFP is a product that minimizes its direct and indirect impact on the environment (including ecological and social aspects) during its whole lifecycle, compared to earlier versions of the same product or similar products that fulfill the same consumer need.

Based on this, the definition of the consumer perspective of EFPs is derived. The main difference is that PEF is perceived and not objectively measured. In the following, when using the term *PEFP*, it represents the consumer perspective. The leading definition of a PEFP applied in this thesis is the following:

A PEFP is a product that consumers perceive to minimize its direct and indirect impact on the environment (including ecological and social aspects), compared to earlier versions of the same product or similar products that fulfill the same consumer need.

Because the judgement of whether a product is eco-friendly or not is made from a consumer perspective and not based on scientific analysis using objective measures (e.g., lifecycle analysis), many ways of communicating the product's EF, e.g., labels, materials, colors, etc., are available to firms and researchers. However, it also means that some products might be described as eco-friendly while not actually minimizing the environmental impact – which is called *greenwashing*. Greenwashing can be described as "the intersection of two firm behaviors: poor environmental performance and positive communication about environmental performance" (Delmas & Burbano, 2011, p. 65). Greenwashing has been addressed by many authors (e.g., De Freitas Netto et al., 2020; Delmas & Burbano, 2011; Fernando et al., 2014; Gosselt et al., 2019; Laufer, 2003; Lyon & Montgomery, 2015; Martínez et al., 2020; T. T. H. Nguyen et al., 2019), but has not been a focal concept of any of the reviewed articles. Thus, it will not be further explicitly analyzed within this thesis.

Given the varying definitions, many operationalizations of PEF and PEFPs can be found in the literature. An *operationalization* is the way how a given concept can be observed and

made measurable (Bühner & Ziegler, 2009). For example, an organic label can indicate EF, thus an organic label could be used as one operationalization of PEF. Section 3.3 provides an overview of the operationalizations of PEFPs used in the reviewed articles. Dangelico and Pontrandolfo (2010; Jäger & Weber, 2020; Larceneux et al., 2012) provide an extensive review of the characteristics which have been used to guide the identification of PEFPs. In addition, most articles included in this thesis have measured the PEF of the stimuli they used. Articles that used *remanufactured / refurbished products* [RRPs] are included in this thesis even without specifically addressing consumers' levels of their PEF. Several studies indicate that consumers do perceive RRPs as beneficial for the environment (Abbey et al., 2015; Atasu et al., 2008; Gaur et al., 2015; Matsumoto et al., 2018; Wallner et al., 2020; P. Wang & Kuah, 2018; Y. Wang et al., 2018). Some studies failed to show the higher PEF, for example, when asked to write "at least three adjectives that you associate with the words 'remanufactured product', students did not mention environmental aspects" (Abbey et al., 2015, p. 492). However, this does not mean that consumers do not perceive RRPs as PEF, it simply shows that other attributes (like used or old) are more strongly associated with it. Due to 1) the increasing attention to RRPs, 2) their crucial part in the circular economy and 3) no conclusive evidence that they are not perceived as eco-friendly products, articles about RRPs are included in the systematic literature review of this thesis.

2.3. Perceived quality (PQ)

The term *perceived quality* consists of two parts (*perceived* and *quality*). Perception has been discussed (section 2.1). For defining *quality*, multiple perspectives and approaches have been taken, e.g., in relation to value, as conformance to given specifications or to consumer expectations (Reeves & Bednar, 1994). This thesis focuses solely on the consumer's perspective. Thus, there are no objective criteria for evaluating PQ, only subjective perceptions. The most widely used definition in consumer behavior literature defines PQ as "the consumer's judgment about a product's overall excellence or superiority" (Zeithaml, 1988, p. 3). She continues that PQ is "(1) different from objective or actual quality, (2) a higher-level abstraction rather than a specific attribute of a product, (3) a global assessment that resembles attitude, and (4) a judgment usually made within a consumer's evoked set" (Zeithaml, 1988, p.3-4). Steenkamp (1990) concluded that all (early) definitions are similar and describe PQ as some form of "fitness for use, given the needs of the consumer" (Steenkamp, 1990, p. 311). Slightly different are Garvin's (1987) eight dimensions of quality (based on operations management): PQ is mainly reputation created from images, advertising, and brand names. The other quality

dimensions are performance, features, reliability, conformance, durability, serviceability, and aesthetics (Garvin, 1987). This thesis follows Aaker's (1991, p. 85) definition of PQ as

"the customer's perception of the overall quality or superiority of a product or service with respect to its intended purpose relative to alternatives".

Aaker's (1991) definition is related to Zeithaml's (1988) definition but extends it with further characteristics regarding the intended purpose. This helps to ensure that PEF and PQ can be clearly separated. PEF is rarely the main intended purpose of a product, because not consuming any product would in most cases be more eco-friendly than using even any PEFP. The intended purpose of food could for example be to nurture the body, to create pleasure or to signal high status. Thus, EF is in this thesis not considered as a dimension of quality itself. However, while not being a quality dimension, PEF could be used as quality cue. Quality cues are "informational stimuli that are, according to the consumer, related to the quality of the product, and can be ascertained by the consumer through the sense prior to consumption" (Steenkamp, 1990, p. 312). Quality cues are an important part of the quality perception process, which starts with acquiring and categorizing cues from the environment, then forming beliefs about quality attributes, and finally integrating these beliefs to the overarching PQ judgement (Steenkamp, 1990). Simply put, quality cues (e.g., country of origin) can be viewed as "what the consumers perceives" and quality attributes (e.g., reliability) can be viewed as "what the consumer wants" (Steenkamp, 1990, p. 313). The meaning and use of the word *attribute* can vary between authors, for example Zeithaml (1988) uses it for rather tangible cues (like color or price), while Steenkamp (1990) uses it for more abstract dimensions (e.g., reliability). Within this thesis, the word attribute is used for both its tangible and abstract meaning because the papers included in this thesis also include both meanings. Figure 4 visualizes the quality perception process based on quality cues as described by Steenkamp (1990).

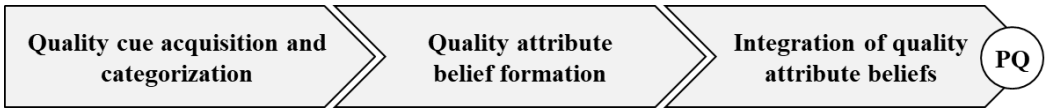


Figure 4: Simplified illustration of the quality perception process.
Source: Based on Steenkamp (1990).

According to McIntyre and Castleberry (1992), a consumer can use multiple cues or a single cue to form the quality perception and even change strategies within the perception process, for example when an intended single cue approach does not allow to differentiate products, other cues can be added to the assessment. Personal (e.g., knowledge) and situational

(e.g., usage goals) factors can influence each stage of the process (Steenkamp, 1990) and are considered within the overarching structure of this thesis as factors influencing the correlations between PEF and PQ (together with product and firm factors).

After explaining PQ and its perception process, closely related concepts will be described in this section to clearly differentiate between them. A distinct concept is *green perceived quality*, which is defined as "the customer's judgment about a brand's (or a product's) overall environmental excellence or superiority" (Y.-S. Chen & Chang, 2013, p. 66). The term, the definition and even the measurement items of green PQ sound very similar to those of perceived quality. However, compared to the clearly distinct definitions of PEFPs and PQ used in this thesis, the concept of green perceived quality is a mixture of these two. This can be seen from the items used for measurement, e.g., "The quality of this product is reliable with respect to environmental consideration" or "The quality of this product is excellent with respect to environmental image" (Y.-S. Chen & Chang, 2013, p. 71). A key word search for only "green perceived quality" within all fields in Web of Science (in July 2021) has yielded only 13 results, showing that the application of this concept is limited in current research. The articles that include green perceived quality (e.g., Y.-S. Chen et al., 2015; Y.-S. Chen & Chang, 2013; Pahlevi & Suhartanto, 2020), typically associate it with similar concepts (like green satisfaction or green trust) that are based on established concepts (like satisfaction or trust) with the tweak of being in relation to environmental expectations or environmental performance. Due to the overlaps with PEF and PQ as well as its limited academic relevance, articles focusing on green perceived quality will not be considered for this thesis.

PQ is closely related to *perceived value* [PV] (V. Kumar & Reinartz, 2016). In their literature review on customer value, Gallarza et al. (2011, p. 183) state that "the value concept is multifaceted and complex" with lacking consent of what it precisely means across consumers, practitioners and researchers. An often used definition of PV is "the consumer's overall assessment of the utility of a product based on perceptions of what is received and what is given" (Zeithaml, 1988, p. 14). Just as Zeithaml's definition of PQ that was established in the same research article is widely used and well respected, her definition of PV is well established. Zeithaml (1988) states that value is based on a tradeoff of give and get components, which is in line with a research stream focusing on this tradeoff (Dodds & Monroe, 1985; Gallarza et al., 2011). PQ is recognized as key part of the "get" side of PV (Gallarza et al., 2011). Thus, scholars agree that PQ is an antecedent of PV (e.g., De Medeiros et al., 2016; Dodds et al., 1991; Macdonald et al., 2016; Zeithaml, 1988). A more recent definition presented by V. Kumar and

Reinartz (2016, p. 37) of PV "as customers' net valuation of the perceived benefits accrued from an offering that is based on the costs they are willing to give up for the needs they are seeking to satisfy". The authors stress that benefits can be diverse, and costs are much more than just the price but for example also learning costs or risk involved. This definition of PV is close to another stream that does not focus on 'the give and get tradeoff', but rather conceptualizes PV as multidimensional construct consisting of dimensions like functional value, emotional value and social value (Koller et al., 2011; Sheth et al., 1991; Sweeney & Soutar, 2001). The dimension called functional value (or performance value), is very similar to quality / performance. For example, an early article defined functional value as "the perceived utility acquired from an alternative's capacity for functional, utilitarian, or physical performance" (Sheth et al., 1991, p. 160). The link to PQ is even more obvious in another definition of functional value as "the utility derived from the perceived quality and expected performance of the product" (Sweeney & Soutar, 2001, p. 211). Thus, research articles that investigate correlations between PEFPs and functional value specifically (defined as perceived quality / performance) are included in the literature review.

In addition, *perceived quality risk* [PQR] (or performance risk) is another concept very close to PQ and will be considered for the literature review (e.g., Brach et al., 2018; Matsumoto et al., 2017). Similar to PV, perceived risk is a multidimensional construct, of which PQR is one dimension besides, e.g., social risk (Stone & Grønhaug, 1993). PQR can be defined as "the possibility of the product malfunctioning and not performing as it is designed and advertised to and therefore failing to deliver the desired benefits" (Matsumoto et al., 2017, p. 969). PQR is detrimental to the superiority of a product, e.g., via reduced reliability or durability. Thus, high PQR indicates low PQ and vice versa (E. J. Lee et al., 2020).

3. Systematic literature review

This chapter builds on the theoretical basis provided in the previous chapter. Via a systematic search and (qualitative) analysis of identified published literature, first answers towards the research questions raised in section 1.2 will be created. This chapter starts by describing the approach how articles included in the review were selected. It then provides insights into descriptive information in terms of publication years, journals, applied research designs and applied measures for PQ. Afterwards, the results about the relationships between PEF and PQ in all kinds of facets are presented before they are discussed and then complemented by suggestions for future research. This chapter closes with a reflection of its limitations and a brief summary.

3.1. Review approach and study overview

The definitions as described above have been used to guide the literature search and assessment, which is explained in this section. The method chosen for this review is a systematic literature review following the guidelines for systematic literature reviews (see Palmatier et al., 2018; Snyder, 2019; J. Webster & Watson, 2002). Using a systematic approach, online databases (APA PsycInfo, EBSCO Business Source Ultimate, and Web of Science) were searched with a set of relevant keywords based on occurrence in high quality literature and fit to the purpose of summarizing correlations between PEF and consumers' PQ. The definitions presented previously have been used to guide the literature search and assessment. The keywords used for systematic literature search were *product* and (*eco-friendl* or green* or sustaina* or environmenta* or ecolog**) and (*consumer or user or buyer*) and (*quality perception or perceive* quality*). Only articles in English and published in journals ranked in the German VHB-JOURQUAL3 were considered to ensure high-quality research. The results were first screened for duplicates and then analyzed according to their content. First, reasons to exclude articles were defined and include the following: Lacking PEF, lacking assessment of PQ, lacking information about relations between PEF and PQ, lacking focus on the consumer or lacking focus on a consumer product. Then, each article was assessed initially based on the title, then on the abstract and where relevant within the full text.

Initial search in March 2020 led to 334 results of articles fulfilling the keyword and journal ranking criteria mentioned above. After excluding duplicated (61) and non-relevant articles (234) based on content analysis as described above, 39 articles remained.

The same search procedures with the additional keyword “ethical” were applied in another database search in August 2021. This led to another 129 search results (exclusive of the 334 results identified previously) identified from the databases. After content analysis, 12 additional articles to be included in this literature review were found.

A final systematic literature search with the same procedures as in 2021 was conducted in the beginning of July 2024 and identified another 346 potential articles published between 2021 and 2024. Of these, 48 were duplicates and 278 were excluded after applying the criteria described previously, leading to 20 articles identified as relevant for the current review. In addition to database search, references and citations in relevant articles were assessed and 17 further articles identified to be included in the systematic review. Overall, this leads to a total of 88 articles summarized within this review.

3.2. Descriptive review results

After the explanation of how articles for the systematic literature review were selected, descriptive results are presented in this section, namely regarding historical development, occurrence in journals, research design applied, and measures of PQ used.

3.2.1. Historical development

Looking at the timeline, the first article in this review was published in 2007. Since 2011, relevant literature was published every year with a peak of nine articles in the years 2019, 2023 and 2024 so far (full year 2024 was not concluded at the time of the final literature search). The number of publications has been growing over the years with a dip in 2020, 2021 and 2022. One possible explanation for this dip might be a shift in researchers’ focus due to the global pandemic caused by the coronavirus disease 2019. Overall, almost all reviewed articles were published in the last 10-15 years, and more than half were published since 2019, showing the recency of the topic in research. Figure 5 shows details of the timeline of publications included in this review.

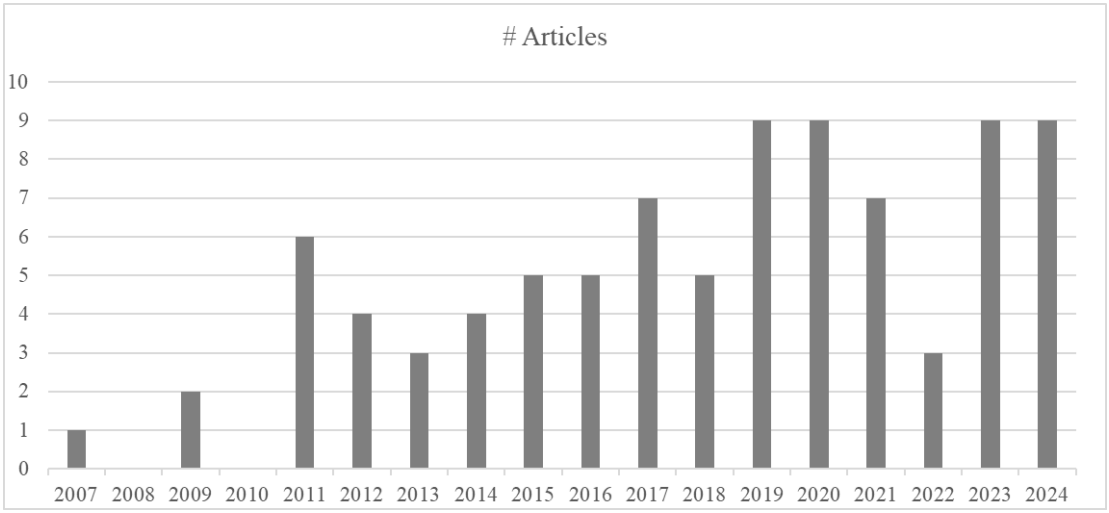


Figure 5: Evolution of publications assessing relationships between PEF and PQ.

3.2.2. Occurrence in journals

By looking at the journals that published the reviewed studies, a better understanding of the spread across disciplines, quality ratings and individual journals can be gained. Most research was published in journals dedicated to sustainability (42 articles) and marketing (29 articles; one article / journal belongs to both areas). Some articles were published in production and logistics (ten articles) or general management (seven articles) journals. 17 journals published more than one article, of which four journals published five or more articles. The Journal of Cleaner Production is the single most active journal (14 articles, 16% of all articles), followed by Sustainability (12 articles, 14%), Journal of Business Ethics (eight articles, 9%) and Journal of Business Research (five articles, 6%). A+ and A journals are represented with 14 articles (16%), B journals with 42 articles (48%), the remaining 32 articles (36%) were published in journals ranked C or lower. Table 3 shows details about the journals.

Table 3: Overview of journals sorted according to number of publications and ranking.

Journal name	Ranking	# Articles
Journal of Consumer Research	A+	1
Journal of Marketing	A+	1
Management Science	A+	1
International Journal of Research In Marketing	A	2
Journal of Industrial Ecology	A	1
Journal of Operations Management	A	3
Journal of Retailing	A	2
Production and Operations Management	A	3
Business Strategy & The Environment	B	4
Energy Policy	B	1
European Management Journal	B	1
International Journal of Production Economics	B	2
Journal of Advertising	B	3
Journal of Business Ethics	B	8
Journal of Business Research	B	5
Journal of Cleaner Production	B	14
Marketing Letters	B	1
Psychology & Marketing	B	3
European Journal of Marketing	C	4
International Journal of Logistics Management	C	2
International Journal of Retail & Distribution Management	C	2
Journal of Consumer Policy	C	1
Journal of Product & Brand Management	C	1
Sustainability	C	12
Sustainable Development	C	1
Journal of Retailing and Consumer Services	C	3
Sustainability Accounting, Management and Policy Journal	C	1
Journal of Consumer Marketing	C/D	4
International Journal of Consumer Studies	D	1

3.2.3. Research design

Methodology and sampling applied in the articles in scope have also been assessed. Some quantitative assessment is used in 85 articles (all but three articles), most often with experiments, but also frequently quantitative surveys. A minority includes qualitative (pre-) studies, while two articles are based on only qualitative research. *Sample sizes* [N] for quantitative studies range from 17 to 1,500 and for qualitative studies from six to 330.

Sampling procedures are important to consider because they can influence the generalizability of results. Most articles by far used convenience sampling, mainly with general consumers, but also quite often with students. Some articles used purposive or theoretical samples, for example only consumers that own a car (Koller et al., 2011; Matsumoto et al., 2017), identify with a green lifestyle (Suki, 2016), or are the main responsible person for shopping in a household (Van Doorn et al., 2020). However, the lines between convenience and purposive sampling can be blurry, e.g., multiple studies categorized as convenience here only considered consumers that are familiar with e.g., green products (e.g., P. M. L. Ng et al., 2024) or recruited participants in supermarkets or malls where the product type of interest is offered (e.g., Forbes et al., 2009; Raj et al., 2023). Other sampling methods included quota sampling, stratified random sampling from a consumer panel or random subsampling from a representative panel. However, this differentiation between sampling procedures (especially for quota / panel samples) can be delicate based on the information provided in the articles. For example, Abbey et al. (2015) state that they used panels, but it remains unclear if they are representative or not (considered here as applying convenience sampling).

Many articles with more than one study used more than one sampling method, for example because they followed different procedures for different surveys / experiments within one article. Regarding the sources from which authors recruited their participants, 28 articles used students as participants in at least one study. 19 articles used participants recruited from Amazon MTurk in at least one study, and some authors used only MTurk for recruitment (e.g., Agrawal et al., 2015; Newman et al., 2014; Nichols et al., 2019, 2023; Techawachirakul et al., 2023). Few authors reported to use other providers like Prolific or Clickworkers (e.g., Abdulla et al., 2024). Overall, (purposive) convenience samples with consumers from the general population remain the most common method.

3.2.4. Measures of PQ

This section addresses the measures used, for example their variety and number of items, to put findings into perspective. Some of the measures applied in the studies within this review seem very different from each other. This means that different studies could be covering different aspects of PQ. However, with PQ being understood as perception of the overall quality or superiority, various ways of capturing this overall impression can lead to the same result. Most (77) articles provide detailed information about their PQ measures. 15 articles use single item measures (Bodur et al., 2016; Bryła, 2019; Can, 2023; Dixon & Mikolon, 2021; Farías, 2020; Forbes et al., 2009; Grolleau et al., 2019; Jäger & Weber, 2020; Larceneux et al., 2012;

Matsumoto et al., 2017; Paparoidamis et al., 2019; Royne et al., 2012; Skard et al., 2021; Techawachirakul et al., 2023; Weisstein et al., 2024) while the others use multiple items: Between two and 16 items, with three items being most common.

The authors of 51 articles included a source for their PQ measures. Six of these articles based their PQ measures on scales from Dodds et al. (1991), three from Larson (1994), three from Newman et al. (2014) - whose articles are part of this review – two from Grewal et al. (1998) and two from Nekmahmud and Fekete-Farkas (2020). The most commonly used measure from Dodds et al. (1991) follows the PQ conceptualization developed by Zeithaml (1988) and takes a consumer-centric view. The second-most commonly used measure by Larson (1994), was developed within a professional buyer-supplier context based on the definition of Garvin (1987) and calls it "measures of supplier product quality" (Larson, 1994, p. 6) so that this measure follows a more technical perspective. The remaining articles reviewed in this work based their measures on varying sources, including for example Agrawal et al. (2015), Larceneux et al. (2012) and Pancer et al. (2017).

Hazen et al. (2017) saw the need for a new measure specific to PQ of RRPs, so they developed and validated a multidimensional measure. The baseline for the development of this scale were Garvin's (1987) dimensions of quality. Starting with seven dimensions (performance, features, reliability, conformance, durability, serviceability, and aesthetics – excl. perceived quality) and expert input for item development, four dimensions remained after factor analysis. These dimensions are 1) lifespan (including durability and reliability), 2) features (including features and aesthetics), 3) performance (including performance and conformance) and 4) serviceability (Hazen et al., 2017). Thus, the scale of Hazen et al. (2017) measures PQ via 16 items (four items in each dimension).

Two recent articles considered explicit (via survey) and implicit (via Implicit Associations Test) measures for PQ. Mai et al. (2019) suggested that the outcomes between implicit and explicit measures may differ. However, the measures used in their article were directed at strong vs. gentle or general positive vs. negative comparisons (especially for implicit perceptions) rather than PQ. Techawachirakul et al. (2023) used both implicit and explicit methods and found the same effect with both methods: Specifically for the implicit measure, participants were both faster and made less errors in the conditions where the PEFP (wine / beer in paper packaging) was paired with negative attributes (e.g., low quality, poor taste, etc.) or

the CP (wine / beer in glass packaging) with positive attributes than vice versa. No other paper was found that considers implicit or unconscious quality perceptions.

3.3. Eco-friendliness operationalization

The underlying theoretical definitions of PEF and PEFPs have been assessed in section 2.2, this section addresses how the reviewed articles have made PEF observable. A variety of PEF operationalizations has been found in the reviewed literature, ranging from general claims (e.g., "better for the environment") via graphical design (e.g., color green), eco-labels (e.g., fair trade, organic, MSC certified or from a fictional standard), or materials (e.g., plastic vs. glass) to forms of the circular economy (e.g., reuse, recycling, remanufacturing). Most articles consider ecological PEF, while the impact of social PEF on PQ is rarely analyzed. To get a better understanding of the focal areas of PEF operationalizations, lifecycle stages are used to group articles (see Figure 6) because lifecycle oriented analysis is common in PEF research (Dangelico & Pontrandolfo, 2010; Sdrolia & Zarotiadis, 2019). Each article in scope is allocated to one or more stages depending on where the eco-friendly impact is happening. For example, organic food is grown using less pesticides and toxins, thus it is considered in the first lifecycle stage because it decreases environmental impact of the raw material. The classification of PEF operationalizations within this thesis is based on the lifecycle stages by Graedel (1995) which are in line with the model displayed on the website of the Life Cycle Initiative (2020) hosted by the "UN environment programme". Other sources do not specifically mention parts manufacturing (Bhander et al., 2003; Halog & Manik, 2011) or do not include transport and / or packaging as separate stage (Bhander et al., 2003; Choi et al., 2006; Nielsen & Wenzel, 2002).

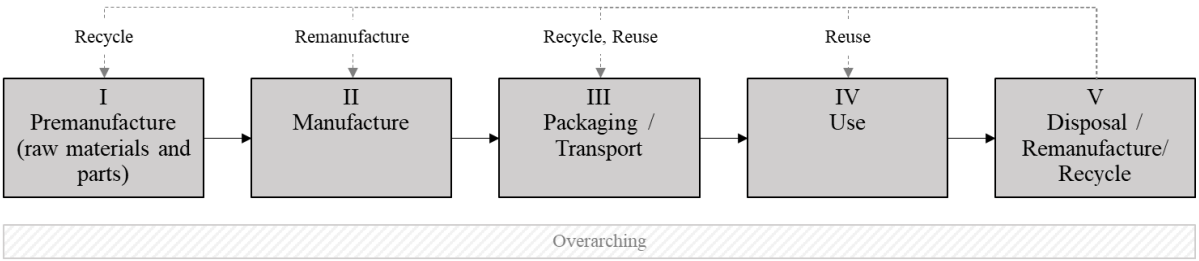


Figure 6: Lifecycle stages including product flow as indicated by arrows.
Notes: Dashed lines indicate circular flows via recycling, remanufacturing, or reuse;
Visualization adapted from Graedel (1995).

Before the allocation of articles in scope to the stages is shown, details about handling specific cases need to be clarified. Recycled, remanufactured, or reused products could be allocated to multiple stages, e.g., in stage V as end-of-life use for the original product or in

either of the stages where the products / materials flow back into (e.g., stage II for remanufactured products). Because the reviewed articles consider the products after the circular flows occurred and not the original products, they are considered in stages I-IV (not in stage V). Products with green design / colors, generic claims (e.g., "better for the environment") or generic eco-labels are not allocated to any stage but shown in a separate category called "overarching" because no environmental impact takes place within any specific stage, or the stage cannot be determined. Nonetheless, a positive effect of green color on PEF has been detected (Pancer et al., 2017; Petersen & Brockhaus, 2017). Table 4 shows the allocation of each reviewed article to the lifecycle stages.

Table 4: Allocation of PEF operationalizations to lifecycle stages by source.
Notes: Multiple allocations possible when multiple products or descriptions are used, e.g., remanufactured, and recycled products or organic claim and green color, etc. (R) indicates that the research is based on recycled, remanufactured, or reused products.

Source	I	II	III	IV	V	Overarching
Abbey et al. (2015)		x (R)				
Abbey et al. (2017)		x (R)				
Abdulla, et al. (2024)		x (R)				
Achabou & Dekhili (2013)	x, x (R)					
Aćimović et al. (2020)	x (R)	x (R)		x (R)		
Agrawal et al. (2015)		x (R)				
Anagnostou et al. (2015)	x					
Atkinson & Rosenthal (2014)	x	x	x			
Bai et al. (2023)	x					
Blanco-Penedo et al. (2021)						x
Bodur et al. (2016)	x					
Borin et al. (2011)	x			x		
Brach et al. (2018)	x					x
Bray et al. (2011)	x	x	x			x
Bryła (2019)	x		x			
Burke et al. (2014)						x
Can (2023)	x					
Chandra Pant et al. (2024)	x					
Y.-N. Cho & Berry (2019)	x	x			x	
Das et al. (2020)				x	x	
Davies & Gutsche (2016)	x	x	x			
De Pelsmacker & Janssens (2007)	x	x	x			
Dixon & Mikolon (2021)						x
Duan & Aloysius (2019)		x (R)				
Elgaaïed-Gambier (2016)			x			

Source	I	II	III	IV	V	Overarching
Farías (2020)	x					
Forbes et al. (2009)	x					
Frank & Brock (2019)			x			
Gleim & Lawson (2014)						x
Gleim et al. (2013)						x
Grolleau et al. (2019)	x			x		
Hazen et al. (2011)	x (R)	x (R)		x (R)		
Hazen et al. (2012)		x (R)				
Hazen et al. (2017)		x (R)				
Herédia-Colaço & Coelho do Vale (2018)	x					x
Hur (2020)				x (R)		
Jäger & Weber (2020)	x					
Kaczorowska et al. (2021)						x
Koller et al. (2011)				x		
Konuk (2018)	x					
Kuah & Wang (2020)	x (R)	x (R)				
Larceneux et al. (2012)	x					
E. J. Lee et al. (2020)	x	x	x			
Lin & Chang (2012)				x	x	x
Lopes, Gomes et al. (2024)						x
Lopes, Pinho et al. (2024)						x
Lou et al. (2022)				x (R)		
Lu & Kwan (2023)	x (R)	x (R)		x (R)		
Lv et al. (2021)		x (R)				
Magnier et al. (2019)	x (R)					
Mai et al. (2019)						x
Margariti (2021)	x					
Matsumoto et al. (2017)		x (R)				
Matsumoto et al. (2018)		x (R)				
Mollenkopf et al. (2022)	x	x				
Monnot et al. (2015)			x			
Mugge et al. (2017)		x (R)				
Munten, & Vanhamme (2023)						x
Newman et al. (2014)						x
P. M. L. Ng et al. (2024)						x
Nichols et al. (2019)	x (R)			x (R)	x	
Nichols et al. (2023)	x	x			x	
Osburg et al. (2024)						x
Ovchinnikov (2011)		x (R)				
Pancer et al. (2017)			x			x
Paparoidamis and Tran (2019)	x			x		

Source	I	II	III	IV	V	Overarching
Petersen & Brockhaus (2017)	x, x (R)					x
Prados-Peña et al. (2023)	x (R)	x				
Raj et al. (2023)						x
Royne et al. (2012)	x			x		x
Ryoo & Kim (2023)		x				
Skard et al. (2021)	x		x			
Steenis et al. (2017)			x			x
Steinhart et al. (2013)						x
Stolz & Bautista (2015)						x
Suki (2016)						x
Techawachirakul et al. (2023)			x			
Tong & Su (2018)	x					
Tymoshchuk et al. (2024)				x (R)		
Usrey et al. (2020)						x
Van de Velde et al. (2009)	x			x		
Van Doorn & Verhoef (2011)	x					
Van Doorn et al. (2020)						x
Waehning & Filieri (2021)	x	x	x			
Wallner et al. (2020)		x (R)				
Y. Wang & Hazen (2016)		x (R)				
Weisstein et al. (2024)	x					
Ye et al. (2022)	x (R)					
Total	***42***	***29***	***14***	***15***	***5***	***27***
Thereof (R)	*10*	*18*	*-*	*7*	*-*	*-*

Looking at the distribution of operationalizations across lifecycle stages, most articles (42) are allocated to the first stage (raw materials and parts) to which for example organic or recycled materials belong. Within stage II, remanufactured products account for 18 out of the 29 articles. Packaging and transport operationalizations (stage III) are represented by 14 articles within this review. Stage IV has been addressed in 15 articles (of which seven articles include reused products) and stage V in five articles. 27 articles use overarching PEF operationalizations, e.g., generic labels, whose impact cannot be allocated to a specific stage. Considering allocation of articles to multiple stages, 27 unique articles consider products that are recycled, remanufactured, or reused across stages I, II and IV.

The focus of existing research lies on the first two lifecycle stages (especially I) and on overall operationalizations. Reasons for this could be manifold, e.g., because consumers typically do not interact with products in the first stages, only in stages IV and V (thus making

their own interaction with the PEFPs irrelevant) or because considering organic products and RRPs are recent trends. Future research can investigate the underlying reasons and implications.

Some operationalizations of PEF have been allocated to multiple stages, but consumers might not be aware or think of all involved stages, e.g., biofuels produce less carbon dioxide [CO2] than conventional fuel during bio-mass production (I) and use (II) (Van de Velde et al., 2009) but consumers might only think about use. Thus, future research can further explore this topic and better understand consumer awareness of PEF impacts. This is important to understand because research indicates that different PEF operationalizations can have different effects on PQ. For example, the centrality of the green claim (e.g., ingredients versus packaging material) affects how consumers' report their PQ (Skard et al., 2021) as well as if the green claim is explicit or implicit in interaction with if the product category is typically green (or not) or if the green attribute is optional (i.e., can be used or not used) or not (Usrey et al., 2020). Also the message valence seems to play a role: When PEF is described negatively (e.g., toxic waste from production affects the environment) it can impact PQ negatively while a positive PEF description does not affect PQ compared to a neutral description (Nichols et al., 2019, 2023). Increasingly stronger negative descriptions do not change PQ beyond the initial impact of a negative message (Nichols et al., 2019).

Similarly, different types of PEF impact vary in their impact on PQ: Bray et al. (2011) found in an early qualitative work, that consumers associated fair trade products with lower PQ but local / free-range food with higher PQ. Another quantitative work found that messages that claim resource use elimination or resource use substitution are related to higher PQ than messages claiming resource use efficiency (Paparoidamis et al., 2019). Petersen and Brockhaus (2017) manipulated color and material and found the effect to depend on product type: Green color is only relevant (negative) for headphones and eco-friendly material (positive) only for garbage bags. Without information about the material, PQ ratings are the same as for CPs, meaning that consumers seem to assume it is not eco-friendly (Petersen & Brockhaus, 2017). Additionally, several studies (Frank & Brock, 2019; E. J. Lee et al., 2020; Pancer et al., 2017) found that combining cues of PEF creates more favorable perceptions than one cue alone. Again, green color can have a negative impact on PQ, and the eco-label too – but PQ remains stable together (Pancer et al., 2017). Differing from Petersen and Brockhaus (2017), consumers assumed that the product would be green without color information (black / white). Frank and Brock (2019) assessed differences between packaging and transport distance. This time though, neither has an effect alone (nor is more impactful), but taken together, PEF is beneficial for PQ.

Combining a sustainability label (MSC certified) with traceability information of the product increases PQ, while using only the sustainability label decreases PQ (E. J. Lee et al., 2020).

3.4. Relationships between PEF and PQ

After having outlined the descriptive results as well as the different PEF opationalizations used in the reviewed articles, this section explains the relationships between PEF and PQ within the reviewed articles. As part of this section, first a summary of direct effects between PEF and PQ is presented, then a brief comparison of effects between industries is given and finally, a high-level overview which constructs have been investigated as impacting PEF – PQ relationships is shown.

3.4.1. Overview and summary of relationships between PEF and PQ

When a product is perceived as eco-friendly, this can have either a positive, a negative or no effect on its PQ. Looking specifically at qualitative research, some positive effects are mentioned (e.g., Burke et al., 2014; Hur, 2020), but negative effects prevail (Abbey et al., 2017; Elgaaïed-Gambier, 2016; Hur, 2020; Ovchinnikov, 2011). Quantitative research presents a multi-faceted picture of the relationships as can be seen across this whole section.

Summary of direct effects of PEF on PQ. Most (54) articles have not reported the direction of direct impact of PEF on PQ (direct impact in this thesis means without considering additional variables and their influence). Twelve articles analyzed the effect and either found none or a specific setting in which there was no significant effect; 20 articles found a positive effect and nine articles found a negative effect. This adds up to 95 rather than 88 because seven articles in total found none and positive effects (four articles) or none and negative effects (three articles). This is possible because the articles used different PEF operationalizations or different samples separately which caused the differences in effect directions. See Table 5 for details.

Table 5: Overview of direct effect of PEF on PQ by source.
Note: 1) means that some indication about the effect is given (further explanation below).

Source	Direct effect of PEF on PQ
Abbey et al. (2015)	Not tested / reported
Abbey et al. (2017)	Not tested / reported
Abdulla, et al. (2024)	Positive
Achabou & Dekhili (2013)	Not tested / reported (anecdotally not linked with poor PQ)
Aćimović et al. (2020)	Negative
Agrawal et al. (2015)	Not tested / reported
Anagnostou et al. (2015)	Not tested / reported

Source	Direct effect of PEF on PQ
Atkinson & Rosenthal (2014)	Not tested / reported
Bai et al. (2023)	Not tested / reported
Blanco-Penedo et al. (2021)	Not tested / reported
Bodur et al. (2016)	Not tested / reported
Borin et al. (2011)	None (compared to neutral setting), Positive (compared to negative eco-impact)
Brach et al. (2018)	Negative
Bray et al. (2011)	Not tested / reported (only qualitative)
Bryła (2019)	Not tested / reported
Burke et al. (2014)	Not tested / reported (only qualitative)
Can (2023)	Positive
Chandra Pant et al. (2024)	Not tested / reported
Y.-N. Cho & Berry (2019)	Positive
Das et al. (2020)	Positive
Davies & Gutsche (2016)	Not tested / reported[1)]
De Pelsmacker & Janssens (2007)	Not tested / reported[1)]
Dixon & Mikolon (2021)	Negative
Duan & Aloysius (2019)	None
Elgaaïed-Gambier (2016)	Not tested / reported (only qualitative)
Farías (2020)	Positive
Forbes et al. (2009)	Neutral / Positive
Frank & Brock (2019)	None (one cue alone) / Positive (two cues combined)
Gleim & Lawson (2014)	Not tested / reported (only qualitative)
Gleim et al. (2013)	Not tested / reported (qualitative and only barriers)
Grolleau et al. (2019)	None
Hazen et al. (2011)	Negative (remanufactured, reused), None (recycled)
Hazen et al. (2012)	Not tested / reported
Hazen et al. (2017)	Not tested / reported
Herédia-Colaço & Coelho do Vale (2018)	Positive (diff between high and low ethical claim)
Hur (2020)	Not tested / reported
Jäger & Weber (2020)	Not tested / reported
Kaczorowska et al. (2021)	Not tested / reported[1)]
Koller et al. (2011)	Positive
Konuk (2018)	Not tested / reported
Kuah & Wang (2020)	Not tested / reported
Larceneux et al. (2012)	Positive
E. J. Lee et al. (2020)	Not tested / reported
Lin & Chang (2012)	Negative
Lopes, Gomes et al. (2024)	Not tested / reported

Source	Direct effect of PEF on PQ
Lopes, Pinho et al. (2024)	Not tested / reported
Lou et al. (2022)	Positive
Lu & Kwan (2023)	Not tested / reported
Lv et al. (2021)	Not tested / reported
Magnier et al. (2019)	Not tested / reported (positive expectation but no comparison with CP)
Mai et al. (2019)	Not tested / reported
Margariti (2021)	Positive
Matsumoto et al. (2017)	Not tested / reported
Matsumoto et al. (2018)	Not tested / reported
Mollenkopf et al. (2022)	Not tested / reported
Monnot et al. (2015)	None
Mugge et al. (2017)	Not tested / reported
Munten, & Vanhamme (2023)	Positive
Newman et al. (2014)	Not tested / reported
P. M. L. Ng et al. (2024)	Positive
Nichols et al. (2019)	Neutral (for positive PEF) / Lower PQ for negative PEF
Nichols et al. (2023)	Neutral (for positive PEF) / Lower PQ for negative PEF
Osburg et al. (2024)	Positive
Ovchinnikov (2011)	Not tested / reported
Pancer et al. (2017)	Negative (one PEF cue), None (two PEF cues combined)
Paparoidamis and Tran (2019)	Not tested / reported
Petersen & Brockhaus (2017)	Not tested / reported
Prados-Peña et al. (2023)	Positive
Raj et al. (2023)	Not tested / reported
Royne et al. (2012)	None
Ryoo & Kim (2023)	Not tested / reported
Skard et al. (2021)	Not tested / reported
Steenis et al. (2017)	Not tested / reported
Steinhart et al. (2013)	Not tested / reported
Stolz & Bautista (2015)	Not tested / reported[1)]
Suki (2016)	Not tested / reported
Techawachirakul et al. (2023)	Not tested / reported
Tong & Su (2018)	None
Tymoshchuk et al. (2024)	Positive
Usrey et al. (2020)	Not tested / reported
Van de Velde et al. (2009)	Not tested / reported
Van Doorn & Verhoef (2011)	(Marginally) Negative
Van Doorn et al. (2020)	Negative (main shopper) / Neutral (students)
Waehning & Filieri (2021)	Positive
Wallner et al. (2020)	Not tested / reported

Source	Direct effect of PEF on PQ
Y. Wang & Hazen (2016)	Not tested / reported
Weisstein et al. (2024)	Not tested / reported
Ye et al. (2022)	Not tested / reported

Several studies provide further insights without applying inferential statistical analyses, for example via descriptive statistics and comparison with scale-midpoints. Out of the articles categorized as "not tested / reported", six articles found tendencies for positive effects (Blanco-Penedo et al., 2021; Davies & Gutsche, 2016; Kaczorowska et al., 2021; Magnier et al., 2019; Mai et al., 2019; Stolz & Bautista, 2015). Further support for no or positive impact were found by Van de Velde et al. (2009): Most (49%) consumers seem to be neutral towards PQ of bio-fuels, while some (38%) belief in its high quality and only a few (13%) disagree with its high quality. Similarly, Achabou and Dekhili (2013) report tendencies of PEFPs not being associated with lower PQ. However, Mai et al. (2019) report negative tendencies when implicitly measured. The earliest article in this review (De Pelsmacker & Janssens, 2007) hints at rather neutral / negative impact of PEF (specifically fair trade) on PQ.

Summary of effects per industry. Studies within 70 articles were allocated to six industries (thereof 11 allocated to multiple industries) based on the products used. 18 articles could not be allocated to any industry either because they used terms like *green products* without industry restrictions or due to limited availability of information. Food and beverages are most often researched and show a rather positive tendency for the influence of PEF on PQ. The direction of influence of PEF on PQ varies a lot in some of the other industries (see Table 6). Table 6 also shows an exemplary (not comprehensive) view of products used in the reviewed articles. These products were used to allocate the studies to industries. Future research could further assess specific industries and compare findings across industries.

Table 6: High-level summary of effects per industry.
Note: Articles can be allocated to multiple industries.

	Food & Beverages	**Clothing, Footwear & Jewelry**	**Personal Care**	**Furnishing & Household**	**Information & Communi-cation**	**Transport**
# Articles	29	9	8	19	14	5
High-level effects	Tendency for positive relations, with neutral relations	Tendency for positive relations, with neutral relations	Highly mixed relations (positive, neutral, and negative)	Highly mixed relations (positive, neutral, and negative)	Tendency for neutral (new products) or mixed relations (RRP)	Mixed relations (positive or negative)
Product examples	Apples, Brownies, Chocolate, Dairy, Fruit, Granola, Jam, Margarine, Pet food, Rice, Salmon, Tomato Soup, Yoghurt; Beer, Coffee, Cola, Milk, Orange juice, Tea, Water, Wine	Luxury Shirts, Running Shoes, Sweater, (Long sleeve) T-Shirt, Watch	Baby shampoo, Bar soap, Body lotion, (Liquid) Body wash, Hand cleaner, Hand sanitizer, Mouth wash, Soap, (Electric) toothbrush	Cups, Coffee machine, Dish-soap, Drain / glass / all purpose cleaner, Garbage bag, Glassware, Detergent, Napkin, Paper, Pen (holder), Toilet paper, (Smart) Vacuum cleaner, Washing machine	Computer, iPhone, Printer, Smartphone, (MP3) Headphones, Headset	Car (parts / engine), Fuel (car)

Summary of constructs impacting PEF – PQ relationships. As shown in the beginning of this section, a high number of articles (54) does not provide inferential statistical analyses about direct effects of PEF on PQ. They focus on the impact of additional variables on the correlations between PEF and PQ. Some of these have a positive impact, for example providing more detailed information about PEFPs leads to higher PQ. For some factors, like product type, or packaging material, the direction of their impact is not that clear and can be either positive or negative depending on their concrete characteristics. Others can have a predominantly negative impact, e.g., a high price discount. Figure 7 shows a summary of the factors identified within existing research and their direction of influence on PQ, clustered into product and firm, and consumer and situational factors.

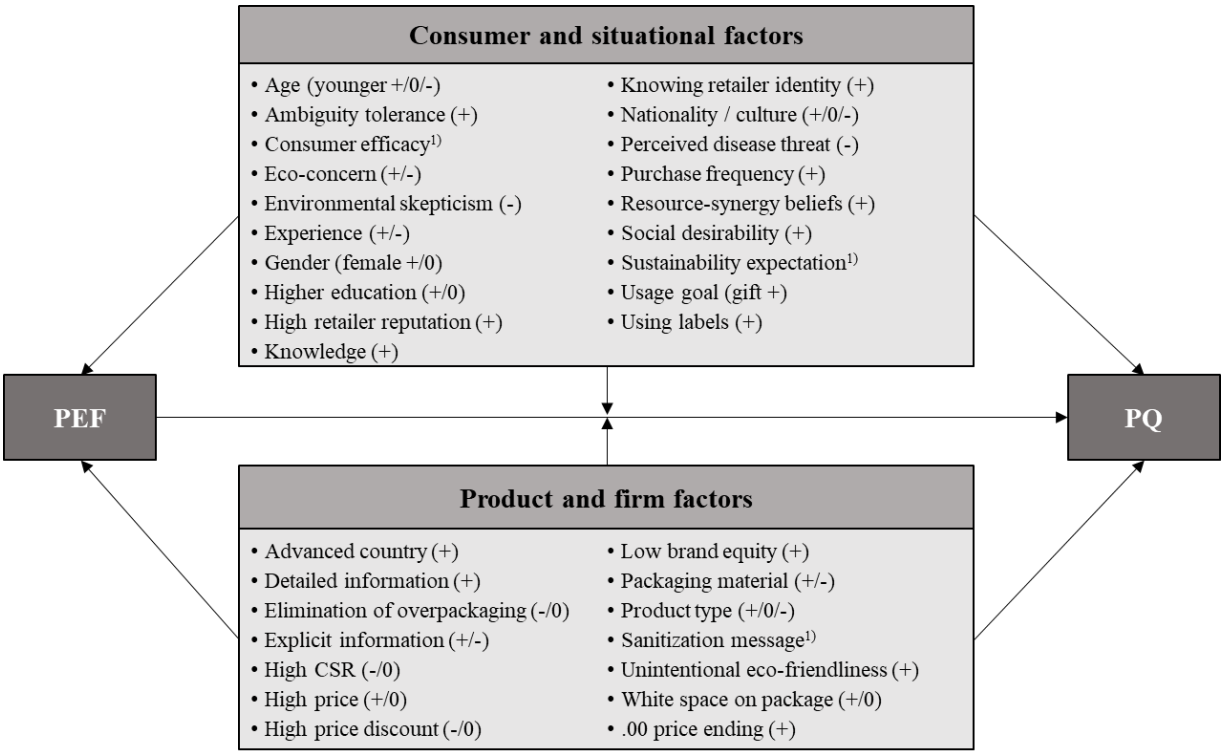


Figure 7: Framework of variables with influence on the relationships between PEF and PQ. Notes: Plus / minus signs indicating simplified direction of influence, e.g., PQ is higher for more detailed information about EFPs. '0' is used to indicate that no impact has been found in at least one study. 1) Factors only analyzed in interaction, thus no influence sign indicated.

Further evaluation and description of the results is clustered into more granular variable categories within the product and firm, and consumer and situational factors. All reviewed articles have been allocated to one or more variable categories depending on the constructs and relationships they investigated. Table 7 shows the variable categories with the number of articles and all sources that include results for each category.

Table 7: Number of articles and respective sources per variable.
Note: Articles can be included in multiple categories.

Variable categories	# Articles	Sources
Consumer and situational factors		
Psychological and behavioral attributes	31	Abbey et al. (2015), Abdulla, et al. (2024), Bai et al. (2023), Bodur et al. (2016), Bryła (2019), Chandra Pant et al. (2024), Y.-N. Cho & Berry (2019), Das et al. (2020), Davies & Gutsche (2016), De Pelsmacker & Janssens (2007), Duan & Aloysius (2019), Farías (2020), Gleim et al. (2013), Hazen et al. (2012), Hur (2020), Jäger & Weber (2020), Koller et al. (2011), Lin & Chang (2012), Lopes, Gomes et al. (2024), Lopes, Pinho et al. (2024), Lu & Kwan (2023), Mai et al. (2019), Mugge et al. (2017), Munten, & Vanhamme (2023), Raj et al. (2023), Royne et al. (2012), Ryoo & Kim (2023), Suki (2016), Van Doorn & Verhoef (2011), Waehning & Filieri (2021), Y. Wang & Hazen (2016)
Demo-graphics	13	Abbey et al. (2017), Abdulla, et al. (2024), Aćimović et al. (2020), Atkinson & Rosenthal (2014), Blanco-Penedo et al. (2021), Hazen et al. (2011), Kaczorowska et al. (2021), Matsumoto et al. (2017), Osburg et al. (2024), Stolz & Bautista (2015), Tong & Su (2018), Van Doorn & Verhoef (2011), Ye et al. (2022)
Usage goals	1	Das et al. (2020)
Retailer reputation and identity	3	Bodur et al. (2016), Konuk (2018), Lu & Kwan (2023)
Product and firm factors		
Type of product	29	Abbey et al. (2015), Abbey et al. (2017), Abdulla, et al. (2024), Achabou & Dekhili (2013), Aćimović et al. (2020), Atkinson & Rosenthal (2014), Borin et al. (2011), Duan & Aloysius (2019), Gleim & Lawson (2014), Hazen et al. (2011), Hazen et al. (2012), Hazen et al. (2017), Herédia-Colaço & Coelho do Vale (2018), Koller et al. (2011), Kuah & Wang (2020), Lu & Kwan (2023), Magnier et al. (2019), Matsumoto et al. (2017), Osburg et al. (2024), Ovchinnikov (2011), Petersen & Brockhaus (2017), Skard et al. (2021), Steinhart et al. (2013), Tymoshchuk et al. (2024), Usrey et al. (2020), Van Doorn & Verhoef (2011), Wallner et al. (2020), Y. Wang & Hazen (2016), Weisstein et al. (2024)

Variable categories	# Articles	Sources
Information	13	Abdulla, et al. (2024), Atkinson & Rosenthal (2014), Brach et al. (2018), Farías (2020), Gleim et al. (2013), Jäger & Weber (2020), E. J. Lee et al. (2020), Lin & Chang (2012), Lu & Kwan (2023), Mollenkopf et al. (2022), Pancer et al. (2017), Skard et al. (2021), Usrey et al. (2020)
Country of origin	1	Matsumoto et al. (2018)
Brand	3	Bodur et al. (2016); Larceneux et al. (2012), Nichols et al. (2023)
Packaging	5	Elgaaïed-Gambier (2016), Monnot et al. (2015), Steenis et al. (2017), Margariti (2021), Techawachirakul et al. (2023)
Price	8	Bodur et al. (2016), Ovchinnikov (2011), Royne et al. (2012), Tong & Su (2018), Ye et al. (2022), Bai et al. (2023), Ryoo & Kim (2023), Chandra Pant et al. (2024)
Firm intentions	1	Newman et al. (2014)
Firm CSR and ethicality	4	Duan & Aloysius (2019); Nichols et al. (2023), Tong & Su (2018), Van Doorn et al. (2020)

After the high-level summaries provided in this section, the next sections (3.4.2 and 3.4.3) will describe each variable individually in more detail. Subsequently, section 3.4.4 will provide information about spillover effects from PEFPs to the PQ of conventional products.

3.4.2. Consumer and situational factors

As mentioned by Steenkamp (1990), consumer and situational factors can influence the quality perception process within each stage of the quality perception process. Identified consumer factors include demographic, psychological and behavioral variables while situational factors include usage goals of the PEFPs and reputation of the retailer who sells the PEFPs. The following sections will further explain the effects of these variables.

3.4.2.1. Consumer psychological and behavioral attributes

Psychological attributes are in this thesis consumer specific characteristics like intelligence, knowledge, emotions, or beliefs. *Behavioral attributes* are in this thesis consumer specific behavioral patterns like purchase frequency, which can result from psychological attributes. Both emotional and rational influences seem to impact consumers' PQ of PEFPs (Davies & Gutsche, 2016). Specifically, interviewees said that because of the good feeling they

get from fair trade products, these products taste better and they also believed in higher standards (ingredients and processes) of these PEFPs (Davies & Gutsche, 2016). Similarly, an overall stronger positive attitude towards PEFPs is associated with higher PQ (Chandra Pant et al., 2024).

Research has found that attributes like environmental (or fair trade) skepticism / concern, concern for social desirability, sustainability expectations, consumer efficacy beliefs, resource-synergy beliefs, purchase frequency, experience / knowledge or ambiguity tolerance further influence the correlations between PEF and PQ. Environmental skepticism negatively influences PQ of PEFPs (Jäger & Weber, 2020; Royne et al., 2012) and environmental / fair trade concern and involvement tend to be positively associated with PQ (De Pelsmacker & Janssens, 2007; Duan & Aloysius, 2019; Farías, 2020; Koller et al., 2011; Lopes, Gomes, et al., 2024; Lopes, Pinho, et al., 2024; Mugge et al., 2017; Van Doorn & Verhoef, 2011), but there are some authors who reported a negative or no correlation with PQ: Specifically, Lin and Chang (2012) found that environmental concern is related to higher amounts of PEFPs used (i.e., mouth wash and glass cleaner), which in turn is related to lower PQ. Similarly, Raj et al. (2023) reported that higher eco-consciousness leads to higher PQR (meaning lower PQ) in their study with Indian consumers. Nichols et al. (2019) found a positive impact of environmental concern for their social operationalization of PEF, but no impact for the environmental operationalization of PEF. Ryoo and Kim (2023) assessed the interaction effect of ethical consumer behavior with price discounts and could not find an impact on PQ. Similarly, Munten and Vanhamme (2023) reported that environmental concern impacts PEF but has no additional significant impact on PQ when accounting for the PEF effect. Comparing the methods applied between authors finding positive versus negative relations, Lin and Chang (2012) and van Doorn and Verhoef (2011) even used the same measure. One potential explanation for the differences in results could be that environmental concern can enhance PQ evaluations in either direction based on early experiences with PEFPs (Duan & Aloysius, 2019; Lin & Chang, 2012), but this tentative explanation needs empirical testing.

Consumers with concern about social desirability (i.e., who want to be seen as responsible and admirable) rate more sustainable products as having higher PQ compared to the ratings of consumers without these concerns (Y.-N. Cho & Berry, 2019). The authors also find the opposite effect for low sustainability products: Consumers without concern for social desirability believe them to have higher PQ than consumers with this concern. Moreover, Y.-N. Cho and Berry (2019) consider the level of sustainability consumers would expect of a

certain product type: Consumers with concern for social desirability always perceive the more sustainable product to have higher PQ, whereas consumers without concern for social desirability only perceive the more sustainable product to have higher PQ when its sustainability level is within their expectations. However, when consumers also have a high efficacy belief (i.e., that their efforts can make a difference), the effect of their sustainability expectations disappears (Y.-N. Cho & Berry, 2019). Thus, concern for social desirability, sustainability expectations and consumer efficacy beliefs all can impact PQ and directly interact with each other. Resource-synergy beliefs (thinking that enhancing eco-friendly attributes goes along with other product improvements) also moderate effects: High resource-synergy beliefs lead to higher PQ (Bodur et al., 2016; Das et al., 2020). Similarly, consumers preferring food from their region, consider it both eco-friendlier and of higher PQ (Bryła, 2019; Waehning & Filieri, 2021).

Consumers' perceptions of specific product attributes like safety, naturalness or nutrition content of organic food are positively related to their quality perceptions (Bai et al., 2023). Similarly, perceived disease threat (i.e., participants' concern about contracting a disease from a product) plays a role for PQ (Lu & Kwan, 2023). Specifically, at low disease threat, participants attribute higher PQ to repurposed products than to new products. Higher disease threat has a negative impact on PQ of repurposed products (Lu & Kwan, 2023). When including a sanitization message to the product description, this negative impact of disease threat disappears – however, the sanitization message decrease PQ at the low disease threat instead of increasing PQ at high disease threat (Lu & Kwan, 2023). The authors provided a tentative explanation that including a sanitization message might remind consumers about a potential hazard.

Purchase patterns play a role for PQ: Consumers who do not or rarely buy or use PEFPs (compared to average users) think their PQ is lower (Hur, 2020; Suki, 2016); When consumers use (organic) labels more during shopping, they associate organic food with higher PQ (Bai et al., 2023). However, these studies cannot confirm causality. Past poor experiences with PEFPs make consumers believe they have lower quality than CPs (Gleim et al., 2013). Thus, early experiences influence PQ which might determine purchase frequency. On the other hand, being more familiar with or knowledgeable about PEFPs can reduce PQR and increase PQ (Abbey et al., 2015; Abdulla et al., 2024; Farías, 2020; Y. Wang & Hazen, 2016), even though Mugge et al. (2017) found the consumer cluster with the highest PQR also to have the highest scores in knowledge and involvement. Overall, a higher purchase frequency could increase knowledge

and thus potentially positively influence PQ. Both explanations could be valid and further explored. In addition, ambiguity tolerance influences PQ positively (Hazen et al., 2012). The authors have used RRPs, but it could be relevant for new PEFPs, especially for consumers unfamiliar with PEFPs.

3.4.2.2. Consumer demographics

Demographics are statistical characteristics of a consumer like gender, age, education, or employment status. The influence of consumer demographics is often considered in research. Gender, age, and nationality / culture can influence the quality assessment of PEFPs. Specifically, women (Tong & Su, 2018; Van Doorn & Verhoef, 2011) and young people (Abbey et al., 2015; Van Doorn & Verhoef, 2011) are found to associate PEFPs with higher PQ in some studies. However, other authors find no differences in demographic control variables like gender, age or education (Atkinson & Rosenthal, 2014; Blanco-Penedo et al., 2021; Ye et al., 2022), especially for reused, remanufactured and recycled products (Abbey et al., 2017; Abdulla et al., 2024; Aćimović et al., 2020; Hazen et al., 2011). A study dedicated to analyze differences in perceptions of sustainable products between age groups, also finds no differing perceptions of overall quality between the groups (Stolz & Bautista, 2015). However, the middle aged group shows tendencies for higher PQ than younger or older people in some dimensions with differences between German and Spanish participants (Stolz & Bautista, 2015) Thus, the influence of demographics like age or education can also depend on other factors, including consumers' nationality (Kaczorowska et al., 2021; Stolz & Bautista, 2015). Specifically, Kaczorowska et al. (2021) compare responses of Polish and Belgian consumers and find that there is no influence of demographic variables for Belgians' responses, while responses of Polish consumers are influenced by age, education and beliefs of their financial situation. Nevertheless, the overall correlation between sustainability and PQ does not differ between the nationalities in their study. However, Osburg et al. (2024) find differences in relationships between PEF and PQ depending on respondents' nationalities: For example, Brazilian consumers assign greater PQ to sustainable luxury products than Indonesian or South African consumers. Matsumoto et al. (2017) find cultural differences between the USA and Japan regarding perceptions of RRPs. They find PQ to be lower in Japan – linked with higher PQR and lower knowledge. These findings could also be linked to differences in psychological attributes, namely skepticism, experience, and knowledge, as further explained in the next section.

3.4.2.3. Usage goals

Usage goals describe in this thesis the intentions a consumer has for a product, e.g., use it for oneself or give it to someone else. Das et al. (2020) find that consumers link higher PQ with PEFPs when they intend to use the product as gift rather than for own consumption. Consumers seem to be looking for different product attributes depending on their usage goals. Further clarifications, e.g., considering routine purchases for others (Das et al., 2020), could shed more light onto the specific effects of usage goals.

3.4.2.4. Retailer reputation and identity

Retailer reputation describes in this thesis the quality associations consumers have with a given retailer, i.e., a store selling products (often from different brands) to consumers. Retailer reputation is known to influence PQ of private label products (Bao et al., 2011). Konuk (2018) shows that this is also true for PEFPs: When consumers have a better image of the store, they rate the PQ of organic private label food higher. Bodur et al. (2016) find that private label PEFPs have higher PQ than CPs only when retailer reputation is high, while no differences occur at low retailer reputation. The authors use private label brands and real retailers that do not differ in perceived familiarity or *corporate social responsibility* [CSR] perceptions. Thus, it remains open if retailer reputation plays a role for (national) brands that are sold across retailers and are not connected to the retailer itself. If that should not be the case, consumers could use retailer reputation to infer brand equity, and retailer reputation would rather be a product and firm factor than a situational factor. Without providing any product descriptions, Sánchez-Gonzaléz et al. (2021) have found that a PEFs' retailer reputation influences product PQ positively. This could hint at retailer reputation being relevant across all kinds of products. However, since no other research addressing retailer reputation in this context was identified, final evaluation depends on future research.

Retailer identity has been operationalized by Lu and Kwan (2023) via the name and photo of a store owner. In their research, knowing the retailer identity increases participants' PQ of refurbished cups but decreases PQ for new cups. Lu and Kwan (2023) assess this relation in the context of disease threat and have hypothesized that introducing the retailer identity reminds participants that an actual person has been in contact also with new products which decreases their PQ.

3.4.3. Product and firm factors

Product (apart from PEF) and firm factors identified in the reviewed articles addressing the correlations between PEF and PQ are described individually. These factors are related to the PEFP itself or the firm producing the PEFP. They include product type, information, country of origin, brand, packaging, price, firm intentions, and firm CSR. Most of these factors can be used as a quality cue by themselves so it can be expected that they interact with PEF as quality cue (e.g., Insch & McBride, 2004; McDaniel & Baker, 1977; H. Park et al., 2020; Rao, 2005; Zeithaml, 1988; Zimmer & Swoboda, 2023). Details are elaborated in following sections.

3.4.3.1. Type of product

In this section, the impact of different types of products on correlations between PEF and PQ is summarized. The term *type* is used in this thesis to refer to different product groups based on their specific characteristics. For example, one type are utilitarian products, another are hedonic products (as parts of the same classification, i.e., utilitarian vs. hedonic), and another type are virtue products, another RRPs and so on. Some of these types are similar, e.g., utilitarian and virtue products, and there could be overlaps, e.g., a remanufactured, utilitarian, slow-moving product. For this review, the product type as stated within the original research has been taken as given and not altered. Several studies suggest that utilitarian, virtue, simple / natural products, gentleness-dependent and products with higher health impact tend to be more positively affected by eco-friendly claims than their counterparts (Borin et al., 2011; Herédia-Colaço & Coelho do Vale, 2018; Skard et al., 2021; Steinhart et al., 2013; Van Doorn & Verhoef, 2011; Weisstein et al., 2024). For gentleness-dependent products, this effect depends however on the centrality of the green attribute (positive for central attribute, negative for peripheral attribute), while PQ of strength-dependent products (e.g., drain cleaner) is negatively influenced by PEF (Skard et al., 2021) for both core and peripheral attributes. The typicality of a product category as PEF or not PEF impacts PQ in interaction with how explicit the claim is (Usrey et al., 2020).

Consumer involvement can be used to classify products: High-involvement products are more expensive and advanced than low-involvement products (Atkinson & Rosenthal, 2014). However, Atkinson and Rosenthal (2014) could not find differences, which could indicate that cost and technological advancement of PEFPs might not be suitable to differentiate PQ of PEFPs. Similarly, Osburg et al. (2024) examine differences between luxury and mass market products (both conventional and sustainable) and find marginal significance for the interaction.

Thus, other forms of involvement (e.g., health impact) might be more relevant for consumers' PQ. Beyond that, the turnover rate of products, i.e., slow-moving vs. fast-moving goods, moderates quality perceptions of PEFPs: Eco-friendly design impacts PQ negatively for slow-moving goods and eco-friendly material impacts PQ positively for fast-moving goods (Petersen & Brockhaus, 2017).

Weisstein et al. (2024) analyze the impact of PEF and product type on PQ of food products via experienced tastiness (participants actually taste the food) and find that PEF increases experienced tastiness for utilitarian products, while the effect for hedonic products is not significant but directionally lower for PEF. Experienced tastiness positively affects PQ.

Not only distinct types of new products are analyzed, but also recycled, remanufactured, reused, and repurposed products. *Recycled* products result from "the process of collecting used products, components, and/or materials from the field, disassembling them (when necessary), separating them into categories of like materials (e.g., specific plastic types, glass, etc.), and processing into recycled products, components, and/or materials" (Beamon, 1999, p. 337). Recycled products are most likely to be perceived as similar to new products (Hazen et al., 2011), which is supported by similar *willingness to pay* [WTP] (Michaud & Llerena, 2011). Magnier et al. (2019) find that consumers expect products made from recycled ocean plastic to be of good quality. Even for luxury fashion, consumers do not link recycling with poor quality; however, they accept packaging to be recycled while opinions about product materials are mixed (Achabou & Dekhili, 2013). Kuah and Wang (2020) do not differentiate between recycled and remanufactured products so that their results cannot be interpreted for recycling specifically. Another study though reports that consumers associate recycling with lower PQ (Aćimović et al., 2020). One explanation for the differing results could be geographic differences (Serbia vs. USA, France, and Netherlands) and respective knowledge / experience of consumers (see section 3.4.2.1 for effects of knowledge / experience).

Next, *remanufactured* products have been assessed by several authors. Remanufacturing means bringing a used product via various process steps into like-new condition (e.g., Hazen et al., 2017; Matsumoto et al., 2017). These products are also called refurbished products (Abbey et al., 2015) and both terms are used as synonyms in this review, being abbreviated as RRP. They considerably help saving resources (e.g., raw materials and energy) and reducing waste (Michaud & Llerena, 2011; S. Webster & Mitra, 2007). Nevertheless, several studies report lower PQ for RRPs than new products (Aćimović et al., 2020; Hazen et al., 2011; Kuah &

Wang, 2020; Matsumoto et al., 2017), while only Duan and Aloysius (2019) do not find differences. Abdulla et al. (2024) do not compare RRPs with new products but find a positive relationship between PEF and PQ of remanufactured products. The two more recent articles reported neutral or even positive relations, thus there could be a shift in consumer perceptions (subject to further research). Recently, Wallner et al. (2020) investigate the impact of design styles on RRPs and found that a simplistic design evokes higher PQ than prototypical design because consumers are reminded of high quality brands known for simplistic design. Similarly, also RRPs in neo-retro design (i.e., design from the past decades for modern technology) have higher PQ because consumers perceive them to be more durable and are reminded of "the good quality of the past" (Wallner et al., 2020, p. 11). Thus, how RRPs look like does influence PQ. The importance of aesthetics is also reported by Abbey et al. (2017), who identify perceived probability of cosmetic and functionality defects as antecedents of PQ. Ovchinnikov (2011) finds that the perception of RRPs is linked to the PQ of the same brand's new products: If a brand has quality issues with new products, PQ for this brand's RRPs is even lower.

Further, *reuse* is "a nondestructive process that finds a second or further use for end-of-first-life solid materials (products or components) without a change of state, excluding melting for metals, plastics and glasses, and pulping for paper" (Cooper & Gutowski, 2017, p. 38) and in this thesis also without bringing the product via industrial remanufacturing into like-new status. Second-hand products, e.g., clothing, are considered as reused products in this thesis. Reuse has an impact on PQ: For example, PEF attributes (hybrid car) can generate a stronger positive effect on PQ for used cars than for new cars (Koller et al., 2011) and PEF and PQ can be positively related for second-hand clothing (Tymoshchuk et al., 2024). Nevertheless, other authors (Aćimović et al., 2020; Hazen et al., 2011) find a negative impact of reuse on PQ.

Finally, a "repurposed product is made from an old product but currently serves a different function" (Lu & Kwan, 2023, p. 3). Lu and Kwan (2023) report relations between being *repurposed* and PQ – however they have not assessed a direct effect in isolation but made their analyses in relation to perceived disease threat. In short, when disease threat is low, repurposed products have higher PQ than new products, but PQ of repurposed products decreases as disease threat increases.

3.4.3.2. Information

Not only does the type of product have an impact, but also the information provided about the product at hand has influence on the effect of PEF on PQ. *Information* describes in

this thesis which and how benefits and product details are communicated. While consumers rate the effectiveness of PEFPs lower compared to CPs in studies by Lin and Chang (2012), the differences disappear when consumers are informed that the product has superior cleaning performance according to scientific tests. However, when Lu and Kwan (2023) added a sanitization message to a refurbished / repurposed product description, PQ decreases if the perceived disease threat is low. They argue that this effect might occur because consumers might only think about potential hazard when receiving the sanitization message. Overall, combining information about PEF and other attributes seems to be important to PQ (Gleim et al., 2013). Gleim et al. (2013), E. J. Lee et al. (2020) and Mollenkopf et al. (2022) report that more detailed information leads to higher PQ of PEFPs, while Farías (2020) find a positive but non-significant effect. More concrete information (e.g., "30% more vitamin C") paired with focus on environmental benefits increases PQ – while general descriptions (e.g., "for more wellbeing") do not (Jäger & Weber, 2020). Mollenkopf et al. (2022) finds that a negative effect of a product recall information is stronger than the positive effect of more transparency on PQ.

The centrality of the green claim also matters: If core attributes (like ingredients) are perceived as EF, the effect can differ from when peripheral attributes (like packaging) are perceived as EF (Skard et al., 2021). Specifically, while a green core benefit decreases PQ for strength-dependent products and increases PQ for gentleness-dependent products, a green peripheral benefit can decrease PQ of both product types (Skard et al., 2021). Skard et al. (2021) also find that a green attribute has lower negative impact on a strength-dependent product when it is peripheral than when it is core. Next to that, the explicitness of PEF claims (i.e., how prominently PEF claims are displayed) can impact PQ (Usrey et al., 2020). For example, when participants learn that a product category is typically green, explicit PEF claims are related to higher PQ, while for a product category described as not typically green, subtle PEF claims are related to higher PQ (Usrey et al., 2020).

Not only detail-level, centrality or explicitness of PEF attributes but also their number can have an influence: Providing two cues of PEF (e.g., color and label, or packaging and mileage, or label and traceability information) increases PQ compared to a single cue (Frank & Brock, 2019; E. J. Lee et al., 2020; Pancer et al., 2017). If an RRP is not only described with text and / or images but participants are actually physically exposed to the product, this can increase its respective PQ (Abdulla et al., 2024). Looking into certifications, a third party certification reduces PQR of organic products, but only when the third party certification is perceived as credible (Brach et al., 2018).

3.4.3.3. Country of origin

The impact of where a product is produced has been addressed in marketing research for decades (Usunier, 2006). Country of origin has been addressed in many studies as factor affecting PQ (e.g., Basfirinci & Cilingir Uk, 2020; Elliott & Cameron, 1994; Insch & McBride, 2004). Previous research has also found that the ecological image of the country of origin can influence consumers' evaluations of PEFPs (Dekhili & Achabou, 2015). One study has looked at PQ evaluations of PEFPs influenced by country of origin: Matsumoto et al. (2018) analyze differences between PQ of RRPs originating from advanced countries (Germany, Japan, USA) or the home country of respondents (Malaysia, Thailand or Vietnam) and find RRPs from advanced countries to have higher PQ.

3.4.3.4. Brand

"A brand is a distinguishing name and/or symbol (…) intended to identify the goods and services of either one seller of a group of sellers and to differentiate those goods and services from those of competitors" (Aaker, 1991, p. 7). It is known that brands can influence PQ of CPs (Kirmani & Rao, 2000). When assessing the influence of a brand on PQ, brand equity is typically used. *Brand equity* comprises assets and liabilities related to a brand that impact the value of a product (Aaker, 1991). Thus, when brand equity is high, the value consumers associate with this brand is high and vice versa. Studies show that PEF is used as quality cue when brand equity is low and that PEF can compensate differences in brand equity (Larceneux et al., 2012). For example, organic food is perceived as higher quality than conventional food for low brand equity products while no differences are found for high brand equity (Larceneux et al., 2012). Similarly, green attributes increase PQ for high priced private label products (lower brand equity) but not for national brand products (Bodur et al., 2016). Brand equity and brand familiarity are connected and familiar brands tend to have higher PQ even though research is somewhat mixed (Nichols et al., 2023). Looking specifically into PQ in the context of PEF, the authors could not find a significant moderation effect of brand familiarity (Nichols et al., 2023).

3.4.3.5. Packaging

Packaging plays an important role in shaping consumer perceptions and is often the first aspect they notice. It is often used by consumers to infer product quality (Zeng & Durif, 2020). Five articles have confirmed the influence of packaging material (Steenis et al., 2017; Techawachirakul et al., 2023), design (Margariti, 2021) and overpackaging (Elgaaïed-Gambier,

2016; Monnot et al., 2015) on PQ. For example, sustainability perceptions of packaging influence overall PQ (Steenis et al., 2017) and perceived packaging functionality (Steenis et al., 2018). When comparing packaging material, "transparent, pot, glass, round, rigid" are associated with higher PQ, while "pouch, bag, carton, flexible" are associated with lower PQ. Similarly, Techawachirakul et al. (2023) find glass packaging of alcoholic beverage to be associated with higher quality than paper packaging – both via survey methodology and via an *implicit association test* [IAT]. Looking at results from lifecycle analysis, glass is the least eco-friendly, while dry carton and pouch packaging score highest (Steenis et al., 2017). Differences between consumer perceptions and lifecycle analysis are confirmed by others (e.g., Boesen et al., 2019; Herbes et al., 2018; Otto et al., 2021), meaning consumers might be easily misled.

Packaging design could also play a role for PQ: Margariti (2021) analyzes the impact of extended "white space" (i.e., space intentionally left empty) for organic food and finds a marginally significant positive effect. In addition, Techawachrakul et al. (2023) give an indication of the relevance of packaging design: While they do not assess PQ itself but rather taste and purchase attention, they find a positive impact of packaging attractiveness on these measures that are related to PQ.

In addition to material and design, the influence of overpackaging (the use of excessive packaging without required functionality, e.g., not for protection, creating unnecessary waste) has been assessed (Elgaaïed-Gambier, 2016; Monnot et al., 2015). Products with overpackaging are perceived as higher or same quality as non-overpackaged products, depending on brand type (Elgaaïed-Gambier, 2016; Monnot et al., 2015). Specifically, PQ of basic private label products is not affected by eliminating overpackaging, thus further investigations how PQ of other products can be kept stable or influenced positively should be made (Monnot et al., 2015).

3.4.3.6. Price

Price describes the amount of money that is asked from a consumer to purchase or use a product. It is an established quality cue (Dekhili & Achabou, 2013; Dodds et al., 1991; Rao, 2005). For CPs, consumers often use higher prices as signals of higher quality (Yang et al., 2019). Impact of price on PQ of PEFPs tends to follow that direction but price does not always play a role for PQ. Recently, Bai et al. (2023) find a positive relation between perceived high prices and PQ for organic food. Similarly, green attributes increase PQ for high priced private label products but decrease PQ for low priced private label products (Bodur et al., 2016). However, in a study by Tong and Su (2018), a high price only induces high PQ without organic

label and with poor firm CSR but showed no effect for PEFPs and / or high CSR. And for upcycled pet food (i.e., made out of food scraps / byproducts of food production) versus conventional pet food, PQ only differs when the product is inexpensive: While there are no PQ differences when the price is high, PQ for upcycled pet food is higher than for the conventional alternative (Ye et al., 2022). Ryoo and Kim (2023) find no effect of price discounts on PQ of fair trade clothing. If consumers believe the price of PEFPs to be overall fair / justified, they associate it with higher quality (Chandra Pant et al., 2024).

Apart from price levels, price endings (.99 vs .00) impact PQ. Without considering price endings, Royne et al. (2012) find no differences, but when the price ending is .00, PQ is higher for PEFPs. Those (.00) endings are often associated with higher perceived price (Schindler, 2006), providing additional support for the positive impact of high prices on PQ of PEFPs. Looking at RRPs, price discount is a quality cue: High discounts (i.e., low prices) are associated with lower PQ (Ovchinnikov, 2011). This could be due to lay beliefs that higher discounts are linked with higher PQR.

3.4.3.7. Firm intentions

Firm intentions describe in this thesis that a firm acts with a specific intention or purpose, e.g., creating an eco-friendlier product. This variable has not been used in PQ research as extensively as brand or price, but Newman et al. (2014) show its relevance: When firms intentionally increase PEF, consumers believe new products to be of lower quality than when firms do so unintentionally (as side effect). Moreover, without information about intentions, consumers seem to assume intentionality and report lower PQ than in the unintentional scenario (Newman et al., 2014). However, Newman et al. (2014) conduct several experiments and have not consistently found this effect of firm intentions on PQ (while effects on PI are consistent).

3.4.3.8. Firm CSR and ethicality

Not only firm intentions are relevant, but also their wider engagement in terms of CSR can be important. According to Turker (2009, pp. 413–414) “CSR can be further defined as corporate behaviors which aim to affect primary social, secondary social, primary nonsocial, and secondary nonsocial stakeholders positively and goes beyond its economic interest”. Nonsocial stakeholders include for example future generations, animals and the natural environment (Turker, 2009). CSR perceptions can influence PQ: For example, research indicates that higher CSR perceptions of a brand lead to higher brand PQ (Ramesh et al., 2019; Van Doorn et al., 2020). However, Ramesh et al. (2019) do not specifically use PEFPs and

neither consider PQ on product level but rather on brand level. Looking at PEFPs, Tong and Su (2018) could not find an interaction effect between firm CSR and product PEF regarding its PQ, while Duan and Aloysius (2019) report lower PQ for refurbished than for new products when CSR is high, but no differences in PQ without CSR information. Van Doorn et al. (2020) find opposing results, i.e., that high firm CSR can mitigate a negative effect of PEF on PQ so that PQ does not differ anymore between the PEFP and the CP when firm CSR is high. However, van Doorn et al. (2020) only find these results with a purposive sample of consumers mainly responsible for household shopping, while there is no effect of PEF and no interaction between PEF and firm CSR for a previous student sample. Explanations for these differing results between authors have not been conclusively discussed in the articles. Several potential explanations seem possible, e.g., using real versus fictitious brands: Tong and Su (2018) use real companies (while Duan and Aloysius (2019) as well as van Doorn et al. (2020) manipulate information provision of fictitious brands), meaning that their high CSR firm (Patagonia vs. Walmart for low CSR) could be so well-known for its PEF efforts that the additional benefit of an eco-label is not relevant anymore. Overall, further research seems required to explain or reconcile the differences in findings regarding the impact of firm CSR.

Nichols et al. (2023) measure perceived ethicality of a firm after negative, neutral or positive messages of their product's manufacturing process and find perceived ethicality of the firm to fully mediate the effect between environmental (or social) information and PQ.

3.4.4. Spillover effects

Apart from assessing the effects and influencing factors of PEF on the PQ of PEFPs, several articles have considered spillover effects of PEFPs on the PQ of CPs. *Spillover effects* occur when "information provided in messages changes beliefs about attributes that are not mentioned in the messages" (Ahluwalia et al., 2001, p. 458). In this case, spillover occurs when the presence of PEFPs does affect the PQ of CPs. Anagnostou et al. (2015) find support for this spillover: When organic coffee is an alternative, PQ of conventional coffee decreases. Similarly, when RRPs from the original manufacturer are available to consumers, their perception of the new products' quality decreases (Agrawal et al., 2015). However, when products are refurbished by independent third parties, their presence increases the PQ of the new product (Agrawal et al., 2015). In line with previous findings, a recent study finds that when perceived similarity between RRPs and CPs is high, PQ of the new product is negatively affected (Lv et al., 2021).

3.5. Discussion of review results and research gaps

After having described the detailed results from the systematic literature review in the previous sections, this sub-chapter provides a discussion by putting results into broader context. It also outlines managerial implications of the results and proposes potential avenues for future research based on the findings described (or lack thereof) and the review's limitations.

3.5.1. Discussion

Reviewed literature shows that PEF is often used as quality indicator – but not always. Most articles have not tested the effect that PEF alone has on PQ but added other variables into the assessment. Different articles found negative, no, or positive correlations. Thus, publication bias, i.e., higher likelihood of significant results to be published compared to non-significant results (Franco et al., 2014), seems not to be a major issue in this area of research. The theories used in the reviewed articles also support a variety of possible correlations. Many theories are not limited to predicting one effect direction, e.g., signaling theory (e.g., Petersen & Brockhaus, 2017), cue utilization (e.g., Herédia-Colaço & Coelho do Vale, 2018), means-end chains (e.g., Hur, 2020), or self-perception theory (e.g., Frank & Brock, 2019). However, some theories are used specifically to explain negative effects, e.g., zero-sum heuristic (e.g., Newman et al., 2014), goal dilution (e.g., Grolleau et al., 2019) and prospect theory (e.g., Y. Wang & Hazen, 2016). The articles in scope suggest first strategies how a negative impact of PEF on PQ might be mitigated, e.g., by providing more specific information (Jäger & Weber, 2020; Lin & Chang, 2012), by activating environmental schemes or using multiple PEF cues (Pancer et al., 2017), or by positioning enhanced PEF attributes as unintentional development (Newman et al., 2014).

To make sure that the PEF operationalization works as intended, many articles have measured if PEF really increased and confirmed successful manipulation. However, no article has explicitly addressed the role greenwashing perceptions might play. Even when consumers perceive EFPs as more eco-friendly than CPs, there might still be doubts about the actual impact or intentions of the firm behind. There is research indicating that PQ of PEFPs with third party certifications / information is significantly better than of PEFPs without third party certification / information (Brach et al., 2018; Delmas & Gergaud, 2021; E. J. Lee et al., 2020). Thus, considering the impact of greenwashing and certifications of PEFPs on PQ could help to further explain their relationships. Many articles (but not all) have explained and defined PQ, but only few have done so for PEF. Creating a better understanding of what PQ and PEF are from a consumers' perspective can help to better address these concepts in future research. Most

studies in this review have focused on the ecological aspect of PEF, with only few considering explicitly social aspects (e.g., fair trade). Since research about the relationships between PEF and PQ has developed relatively recently, it can be expected that more research will be conducted which can then benefit from enhanced conceptualizations and further clarify the correlations between PEF and PQ.

Existing research has been mainly created in the past decade with more publications in the second half than in the first. Research is published in journals from several fields, yet focusing on sustainability and marketing. Measures used are quite different between studies, meaning that direct comparability may be partially limited, which could be related to some variations in findings. The need for a new measure is further outlined as research gap later in this chapter. Looking at PEF operationalizations, most articles have signaled PEF via organic labels, information about products being remanufactured / recycled, or verbal claims (e.g., better for the environment). From a lifecycle perspective, the first two stages (raw materials / parts and manufacturing) are most prominent, whereby effects in the second stage are almost exclusively via remanufacturing and not e.g., processes with lower energy use or pollution during original manufacturing. From an industry perspective, food and beverages are the most researched products and show a tendency for more positive correlations than products in other industries. Additionally, congruency of product attributes, e.g., color and eco-label (Pancer et al., 2017) or natural product and eco-claims (Herédia-Colaço & Coelho do Vale, 2018), influences PQ positively, which might be an interesting area for practitioners to explore. One potential explanation for this effect could be that congruent attributes lead to higher credibility and trust by consumers. Also the centrality of green attributes (e.g., ingredients versus packaging) can influence how PQ is affected (Skard et al., 2021).

PEF affects products differently according to their type – with a tendency for higher PQ of utilitarian, natural and fast-moving goods. However, some of the product examples used in the literature could be allocated to another type than what they were meant to represent (e.g., garbage bags as utilitarian products rather than slow-moving products) so that investigation of different product type classifications can further clarify product type implications. As specific product type, remanufacturing is not only quite prominent in research, but RRPs tend to be perceived as lower in quality than new products. Considering the relevance of strength of a product to differentiate between strength- and gentleness-dependent products (e.g., drain cleaner as strength-dependent and body lotion as gentleness-dependent), there is evidence that

PEF has a negative effect on PQ and product preference for products that are expected to be strong (Luchs et al., 2010; Mai et al., 2019; Skard et al., 2021).

Other variables known to influence PQ (e.g., brand, price, or firm CSR) can influence or cancel out the impact of PEF. Several articles have analyzed their impact and came to partially contradicting results. These differences could be caused by a variety of reasons e.g., other underlying factors not analyzed, differing belief / disbelief in the PEF operationalization, etc. For example, while perceived eco-friendly packaging material is accepted, consumers seem to misjudge the actual EF of packaging (Steenis et al., 2017). Thus, further education and knowledge building could help consumers make more accurate judgements and behave more eco-friendly. Nevertheless, some "old-fashioned" associations seem to have remained so far, e.g., that some consumer groups believe overpackaging to show higher PQ (Elgaaïed-Gambier, 2016). In addition, consumer demographics may play a key role, which needs to be better understood. The impacts of e.g., gender and age differ between articles. Several other characteristics influence PQ of PEFPs, e.g., higher ambiguity tolerance (Hazen et al., 2012). Nevertheless, also within psychological and behavioral attributes, some inconsistencies and open questions remain to be addressed in future research. Finally, situational circumstances like usage goals and retailer reputation seem to affect PQ evaluations of PEFPs, but further assessment is needed.

3.5.2. Practical implications

Faced with increasing pressure from public bodies and consumers to increase the share of eco-friendly products, firms can draw on insights from this review to design and market their products to reach high PQ. This in turn can create a positive impact on WTP, PI and actual purchase (De Medeiros et al., 2016; Homburg et al., 2005; Mai et al., 2019; Tong & Su, 2018). Based on the identified product, firm, consumer and situational factors, firms can adjust their marketing strategies. Marketing campaigns should address PEF of products and combine it where necessary with other quality cues that were summarized in this review. For example, when preparing a new marketing initiative, managers can make use of the existing insights and, for example, provide detailed information about PEF and product quality, while informing consumers about their intentions to create a better performing product. Pricing, packaging and maybe even targeting strategies for specific customer groups can be developed to maximize PQ. Combining these insights with the knowledge about which products tend to be seen as of lower PQ by consumers, marketing managers can focus their efforts on specific products and innovation managers can purposefully address them in product enhancements. In addition,

firms can use established measures to understand the perceptions consumers have of their PEFPs better, allowing them to identify any issues early on. This will allow research & development or innovation departments but also production and supply chain to address these topics and develop products that are perceived as better by consumers. The identified consumer and situational factors can also be used by marketing managers to increase PQ of their products, for example by targeting specific consumer segments or distribution channels and partners. Future research on the identified research gaps will continue to provide insights that managers and firms can use to develop and market PEFPs that are well received by consumers.

3.5.3. Research gaps

Based on the literature review, research gaps have been identified, which can be grouped into three main areas: 1) Detailing the underlying concepts PEF and PQ (section 3.5.3.1); 2) Enhancing empirical evidence (section 3.5.3.2); 3) Assessing new variables that might impact the correlations between PEF and PQ (section 3.5.3.3). Priority should be given to the conceptualization and deeper understanding of the impact of PEF operationalizations as baseline for following research.

3.5.3.1. Detailing the underlying concepts PEF and PQ

As shown in this review, PEF can be operationalized in many ways (see section 3.3). First studies showed that different operationalizations can lead to different PQ evaluations. Thus, consumers' perspectives of what PEF is, which operationalizations have more influence on PQ and why these differences exist, remain mostly unanswered in the current research. Further exploring these questions can lead to a better understanding of PEF and PQ. When doing so, using the lifecycle stages might be further beneficial in explaining differences found in this review, e.g., between recycling and remanufacturing, and bridging the gap between scientific PEF assessment and consumer perspectives. For example, a potential explanation for a more positive perception of recycling compared to remanufacturing or reuse could be that recycling flows back into the initial lifecycle stage, thus further away from the consumer.

In addition, understanding if the presented working definitions and distinction between PEF and PQ are supported from the consumers' perspective is important for future research. Specifically, this paper presented PEF as potential quality cue rather than quality dimension, which should be empirically analyzed. In addition, consumers might use PEF as quality cue only under specific circumstances. Determining these circumstances in future research can be helpful both to researchers and practitioners. Qualitative research could dig deeper into the

previously mentioned topics and answer fundamental questions where later studies can then be built upon. Understanding differences between various forms of PEFPs (e.g., organic, refurbished, etc.), including forms that have not been assessed in relation to PQ could be explored. For example, considering different labels, including new types like carbon labels (e.g., Wong et al., 2020) could increase knowledge in this area. New operationalizations could be investigated, e.g., in the comparingly well researched food industry. This includes sustainable production methods like Community Supported Agriculture (e.g., Samoggia et al., 2019) and vertical farming (e.g., Touliatos et al., 2016) or new products like value-added surplus products (e.g., Bhatt et al., 2018) and insects (e.g., Baker et al., 2016).

3.5.3.2. Enhancing empirical evidence

This section describes ways to deepen the understanding of identified correlations between PEF and PQ. To add a new perspective and gain further support for the identified correlations and effects, methods of data collection could be extended further, e.g., by considering social media posts. This method has been applied to investigate PEF implications (e.g., Blasi et al., 2020) but not in relation to PQ. Other methods of data analysis like *fuzzy-set qualitative comparative analysis* [fsQCA] could shed more light on this topic, e.g., by understanding possible configurations for high PQ of PEFPs. A recent study by Scarpi et al. (2021) has used this method to understand under which configurations consumers intend to switch to bioplastic products but they have not looked at PQ separately.

Still from a methodological view, this review has shown that numerous different measures are used in the presented research, which can reduce comparability between findings. Thus, a new measure (based on a thorough conceptualization) that can be applied both to PEFPs and CPs could aid future research. Apart from Hazen et al. (2017), only two articles creating dedicated PQ measures within the past 10 years have been found in online databases (Das Guru & Paulssen, 2020; Molina-Castillo et al., 2013). However, neither has been reportedly designed, tested, or validated in the context of PEFPs. Molina-Castillo et al. (2013) created a comprehensive index using formative and reflective items to measure overall product quality. They are focusing though on the managers' perspective and objective measures rather than consumers' perceptions. While the measure of Das Guru and Paulssen (2020) is focused on the consumer, it is designed to measure experienced quality, thus potentially not applicable before actual product use. It is relatively long (24 items), which can make it hard to combine with other constructs without losing participants' attention or motivation. Thus, a new measure to capture consumers' quality perceptions of PEFPs addressing the mentioned drawbacks and potential

other ones would be beneficial. In addition, developing and applying a measure for capturing implicit PQ (e.g., unconscious reaction measures via reaction time to certain stimuli) is another interesting path to deepening the knowledge about PQ. Reasons why explicit and implicit measures could differ are diverse, e.g., because consumers might want to nurture their self-esteem or behave in a socially acceptable manner. Implicit measurements of PQ link back to the definition of perception, which for this thesis includes unconscious effects (see section 2.1). Other authors might argue that only conscious effects should be measured and thus not see the need or applicability of implicit measures.

Considering the amount of research available and the differences between product types, further research could investigate if any and then which product categorization is most suitable to account for PQ of PEFPs. Additionally, extending evidence for variables that have been analyzed in one or very few articles only (e.g., firm intentions, usage goals or retailer reputation) or that led to highly differing results across studies (e.g., firm CSR) can fortify and develop the overall knowledge. As specific examples, for usage goals variations like gift-giving vs. routing purchases for others, and for retailer reputation the influence on branded / national brand PEFPs (compared to private label PEFPs) could be assessed. Finally, while a few studies (Kuah & Wang, 2020; Matsumoto et al., 2017; Osburg et al., 2024) have gathered and compared cross-country data, most research has not (or did not disclose). Gaining a better understanding of cultural differences could uncover new insights and implications.

3.5.3.3. Assessing new variables

As the review has shown, multiple variables influencing the correlations between PEF and PQ have been identified, but the interplay with additional variables across product, firm, consumer, and situational factors can be further explored. Their effects on the correlations between PEF and PQ do not have to be tested simultaneously in one overarching model but can be analyzed one by one or in small groups of variables just like previous studies have done. Considering all variables at the same time would not only be a huge task for researchers and participants, but also not suitable for all products, like warranty for food products. Identified variables for future research are summarized in Figure 8 and introduced in the following section. The list of variables is extensive but not exhaustive. It has been developed based on the author's assessment of current literature (within this review and beyond) with the intention of providing starting points for future research.

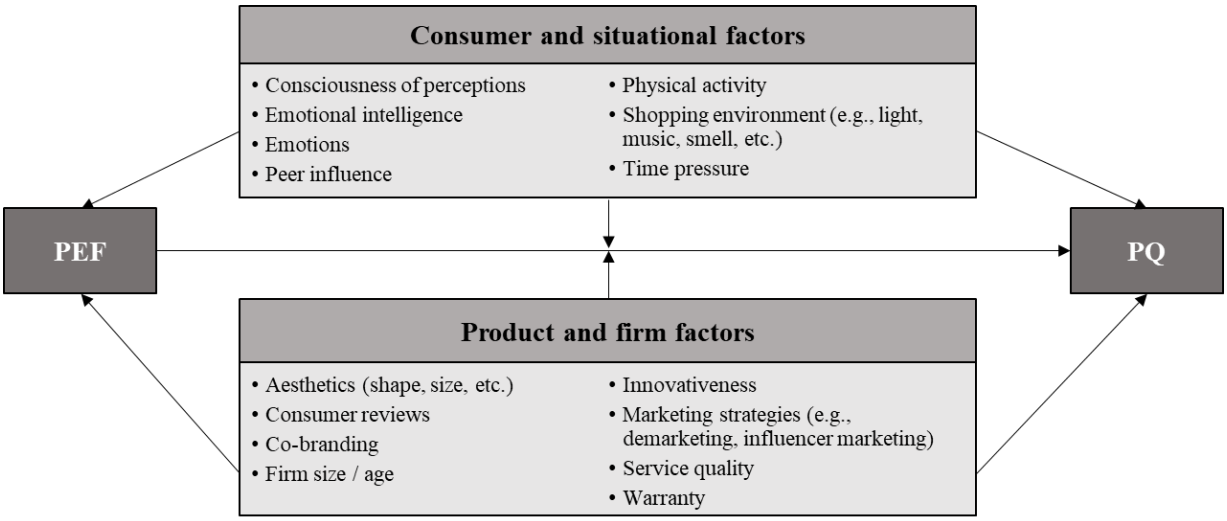


Figure 8: Proposed constructs for future research.

Additional product and firm factors to consider include warranty, innovativeness, aesthetics, co-branding, and reviews (product factors), as well as marketing strategies, service quality and firm size / age (firm factors). By giving consumers suitable warranty policies, PQR can be reduced and PQ potentially increased (Matsumoto et al., 2017; Ovchinnikov, 2011; Y. Wang et al., 2018). Given that warranties are legally binding, they might influence PQ more than less costly quality signals (Frota Neto et al., 2016). In addition, product innovativeness could be linked to PQ. Higher perceived innovativeness leads to higher perceived PEF and ultimately to higher adoption intentions (Paparoidamis et al., 2019). Given the link between PI and PQ, higher PQ could also be related to perceived innovativeness. Looking at product aesthetics, e.g., color (e.g., Hagtvedt & Brasel, 2017; Pancer et al., 2017; Sundar & Kellaris, 2017), shape (Normann et al., 2019; Rahinel & Nelson, 2016), and texture (Karana, 2012) of products, packaging or labels influence perceptions. Thus, testing the effects of aesthetics on PQ could assist in product design. Another area is the impact of co-branding, where the brands of constituent firms influence PQ (Besharat, 2010). While the impact of a single brand has been assessed, combining brands could e.g., follow that direction (maybe even make PEF irrelevant) or create different results. Additionally, product related customer reviews are often connected with quality and play a role in the PEF context (Fernando et al., 2014), but have not been assessed in relation to their influence on PQ.

At firm level, understanding the influence of factors like marketing strategies, service quality and firm size could provide further insights into the correlations between PEF and PQ. For example, the influence of demarketing could be worth assessing. Demarketing means that consumers are discouraged to buy products often for environmental or health reasons (Grinstein

& Nisan, 2009) and does influence perceptions (Sekhon & Armstrong Soule, 2020). For PEFPs, it could either result in increased PQ because it is costly and products should perform better (Miklós-Thal & Zhang, 2013) or decreased PQ because products could be worse when the company counts only on one time sales. Another marketing strategy to investigate is influencer marketing which has a differentiated impact between age groups on sustainable purchasing (Johnstone & Lindh, 2018). Influencers can be seen as having an own brand making these analyses related to co-branding research (Kupfer et al., 2018). Additionally, service quality (e.g., consulting, after-sales, etc.; including e-service quality) could be another factor influencing PQ of PEFPs. E-service quality is positively related with PI (Ahmad & Zhang, 2020). Finally, firm size and / or age could be relevant, e.g., the PEF efforts of small / young firms could be more credible leading to higher PQ, or the resource limitations could be perceived as stronger leading to lower PQ.

Within consumer and situational factors, emotions, and consciousness (consumer factors), as well as shopping environment, peer influence, and time pressure (situational factors) could be explored. Research has shown that emotions influence PQ (Antonetti et al., 2019) and sustainable purchasing behaviors (Antonetti & Maklan, 2014a, 2014b) but the connection between different emotions, PEF and PQ has not been explicitly made. For example, consumers with a feeling of disgust might wrongly attribute this feeling to the PEFP and experience lower PQ. In fact, both Abbey et al. (2017) and Abdulla et al. (2024) find a negative correlation between perceived disgust and PQ. At the same time, fear and disgust seem to be not significant as mediators next to PQ (or as mediator) between PEF and PI (Abdulla et al., 2024; Lu & Kwan, 2023). Thus, overall further assessment of the role of emotions is needed. Additionally, not only emotions but also emotional intelligence might be an interesting avenue for research on PEF and PQ. Emotional intelligence has been shown to impact PI of sustainable products (Sarkar et al., 2022) and might influence PQ. In addition, the role of consciousness in quality perceptions of PEFPs could be further explored. First findings are provided by Mai at al. (2019), yet open questions remain. Apart from applying measures for unconscious evaluations, understanding the reasons and consequences of potential differences between conscious and unconscious PQ could lead to important theoretical and practical contributions. The need for measures to do so has been highlighted at the beginning of this section.

As situational factors, the shopping environment and specifically scent and music (Bosmans, 2006; Chebat & Michon, 2003; Mattila & Wirtz, 2001; Michon et al., 2005), visual warmth (Baek et al., 2018) or space allocation (Sevilla & Townsend, 2016) have been shown

to influence evaluations. The influence of these variables on PQ of PEFPs could differ from CPs, e.g., because attribute congruency plays an important role. Thus, even more "natural" music could have a positive impact on PQ. Also, differentiating between physical and virtual environments might yield new insights, e.g., about the relative importance of attributes (Lombart et al., 2020). Peer influence has not yet been assessed in this context, so that understanding the effect of peer behavior on PQ might create new insights. For example, consumers with peers present that have high PQ for PEFPs might also show higher PQ themselves (and vice versa). Another factor, time pressure, has been shown to influence PQ (Suri, 1996) and can change which information receives most attention (Steenkamp, 1990). Again, it could have different effects for PEFPs than for CPs. In addition, physical activity (of the consumer) might be interesting to investigate further: Sarkar et al. (2022) have shown that active people rather choose a sustainable product than non-active people. The role of PQ has not yet been analyzed within this context but could further increase understanding of the overall topic.

3.5.4. Limitations of the literature review

This literature review has some limitations. First, based on the selected approach and search criteria, relevant articles might not have been found and could thus be missing in this review. The list of keywords for the literature search was on purpose relatively extensive to reduce the probability of overlooking relevant research but nevertheless, some studies might not have been identified. Similarly, limiting the scope to journals that are ranked within the German VHB-JOURQUAL3 list ensures high quality research but might overlook some interesting findings in other journals, books, or publications. Second, the exclusion of articles after screening of the abstracts involved some subjectivity. This subjectivity was limited by conducting multiple screenings over time and following defined criteria for selection. Third, the focus of this review was deliberately put on PQ, but similar analyses with other concepts such as PV, PI or actual purchasing might yield additional insights.

3.6. Summary of literature review

This systematic review chapter has summarized and synthesized high-quality literature mainly from the fields of sustainability, marketing, production / logistics, and management. Based on 88 articles published in journals with VHB ranking of D or better, the correlations between PEF and PQ have been assessed and clustered. Even though PQ is not a new concept in consumer studies, studies about the specific connection with PEF have emerged over the course of the last decade. The literature shows that quality perceptions of PEFPs are often but

not always different from CPs, that their evaluation interacts with other product, firm, consumer, and situational factors, and that there are still inconsistencies and gaps to be further explored. Based on existing findings, this chapter has identified key areas for future research. Those topics will likely provide new insights for both researchers and practitioners.

The findings from this systematic literature review will further inform the following empirical research within this thesis. In the next chapter, the qualitative research element of this thesis is introduced and described.

4. Qualitative research

After the previous literature review has outlined key findings and gaps in published literature focusing on relationships between PEF and PQ, the next step in this thesis is to extend the existing literature, which is intended first via a qualitative primary research and then followed by quantitative research (see chapter 5). The systematic literature review has allowed to define and sharpen the research questions for the primary research part, and to create a thorough understanding of the broad body of research – even from articles not included in this research that were part of the research process.

This chapter is structured as follows: In the next section, the aim and research questions of the qualitative research within this thesis are laid out. Thereafter, the research method, including data collection and analysis as well as sampling is introduced in detail. Following, results of this research are presented: PEF of consumer products is conceptualized and relationships with PQ explored and synthesized. A discussion with theoretical and managerial implications as well as limitations concludes this chapter.

4.1. Aim and research questions

To further the understanding of relationships between PEF and PQ and their interaction with additional constructs, this research adopts a qualitative approach to capture the nuances and try to uncover new ideas and insights. The aim of this qualitative research is twofold: First, to develop a definition and conceptualization of PEF specifically from the consumer perspective. When conceptualizing PEF, consumer understanding of PEF as unidimensional (i.e., referring to one single dimension) or multidimensional (i.e., referring to multiple distinctive yet interrelated dimensions) construct will be investigated. Second, to identify consumer beliefs about cor relationships relations between PEF and PQ via assessing their self-reported reasons for those relationships (instead of abstract theoretical considerations). Specifically, it intends to explore reasons for (positive, negative, or no) impact of a product's PEF on its PQ that consumers can consciously report on.

Understanding the consumer definition and conceptualization of PEFPs as well as the correlations between PEF with PQ in more detail offers several contributions: First, it can support researchers by providing a basis for empirical and quantitative research, in designing future studies that consider the developed concepts, and to better put seemingly contradictory findings to date into context. Second, it can inform and assist policymakers in aligning perceptions with actual EF, e.g., through additional education or adjusted regulations. Finally,

it can also support managers in understanding and addressing consumer perceptions, thus enabling them to effectively design and market their products.

The scope of this qualitative research is limited to *adult consumers*, i.e., people aged 18 or older that regularly purchase tangible consumer products (regardless of the products' EF or PEF). This age limit is introduced for multiple reasons, including: First, ensuring consistency with previous research as those samples were typically focused on adults. Second, limiting the potential influence of differences in cognitive capabilities between children / young adolescents and adults (Icenogle & Cauffman, 2021) on research results. Third, following ethical considerations regarding the ability to give informed consent to participate in the research.

To achieve the aim, a theory about PEF and PQ is developed using the *grounded theory* [GT] approach. A theory can be called "a statement of concepts and their interrelationships that shows how and/or why a phenomenon occurs" (Corley & Gioia, 2011, p. 12). Such a statement of concepts and their interrelationships allows deepening the knowledge about the correlations between PEF and PQ. Existing literature that links PEF with PQ has drawn from various theories, including e.g., signaling theory (e.g., Petersen & Brockhaus, 2017), cue utilization (e.g., Herédia-Colaço & Coelho do Vale, 2018), self-perception theory (e.g., Frank & Brock, 2019) , zero-sum heuristic (e.g., Newman et al., 2014), goal dilution (e.g., Grolleau et al., 2019) or prospect theory (e.g., Y. Wang & Hazen, 2016). However, while some of these theories can be combined in the context of PEF and PQ, other theories let researchers derive different predictions about the correlations between PEF and PQ. Additionally, none of the theories mentioned above were tailored specifically to the context of PEF. The current research is focused on the concepts of PEF and PQ and their specific interrelationships. By engaging directly with consumers and their perspectives, this research seeks to further uncover dimensions and factors shaping these perceptions, thereby contributing to a more holistic understanding of PEF in consumer products.

Derived from the aim of the qualitative study within this thesis, the following research questions guide the work:

1. How do adult consumers conceptualize and define perceived eco-friendly products (PEFPs)?

2.a) Do adult consumers perceive eco-friendliness (EF) in consumer products as unidimensional or multidimensional construct?

2.b) If perceived eco-friendliness (PEF) is multidimensional, what are the underlying dimensions in the context of tangible consumer products?

3.a) How do adult consumers perceive relationships between perceived eco-friendliness (PEF) and perceived quality (PQ) across different product types / industries?

3.b) Which constructs influence the perceived relationships between perceived eco-friendliness (PEF) and perceived quality (PQ) in consumer products?

4.2. Qualitative research methods

In this methodological part, the choice for qualitative research and specifically GT as applied research method is explained first. Next, the GT approach and its application within the data collection, preparation and analysis (including coding, categorizing, theory building and review) of this thesis are described in detail, following the guidance of Charmaz (2014). The final sample is introduced before the results are then presented in section 4.3.

4.2.1. Choice of research method

The choice of research method is derived closely from the research questions presented in the previous section and from the existing insights of the status of research in this field informed by the previously conducted systematic literature review. Thus, a qualitative approach was chosen. The key reasons for this choice are explained in the following:

First, a qualitative approach allows to extract and understand nuanced subjective perceptions and beliefs (Cooper & Schindler, 2014): This is needed to address the research questions revolving around subjective phenomena like PEF or PQ. The potential influencing factors on the relations (research question 3.b) may be subjective phenomena like specific beliefs or previous personal experiences, highlighting the necessity of exploring the views of individual participants in depth. In all research questions, the perspective of consumers is in focus. Applying a qualitative interview-based method with consumers for this research allows to keep that focus on consumer perspectives throughout the entire process.

Second, theoretical development is still evolving within the context of PEF and PQ, suggesting value in using a qualitative approach (Creswell, 2009). Qualitative research allows to generate new frameworks and theories emerging from the data and provides the flexibility to explore patterns, relationships or new factors that may yield new contributions. For example, using a qualitative approach allows to develop potential dimensions of PEF bottom-up and

holistically rather than starting with pre-conceived hypotheses that may overlook specific aspects. Thus, the qualitative approach suits the explorative nature of the research questions best.

Third, qualitative methods allow to gather and analyze rich data, meaning data that are in depth, detailed, and containing a lot of information and context. Conceptualizing complex constructs like PEF and its potential dimensionality, requires the usage of rich data – especially considering the intended application to different product categories. The aim to understand potential dimensions of PEF and relations to PQ across different product categories highlight the need for a thorough contextual understanding which can be gathered in qualitative research.

Within qualitative research, several different methodologies exist, all with specific methods, goals, and applicability. Creswell (2008) describes five key methodologies for qualitative research: Narratives, phenomenology, ethnography, case studies and GT (see Table 8). GT is the chosen methodology for this thesis. The GT methodology was initially established by Glaser & Strauss (1967). It guides the researcher to immerse themselves in the data and then reach a higher, more abstract level, creating a theory grounded in the data. In other words, the "procedures of grounded theory are designed to develop a well-integrated set of concepts that provide a thorough theoretical explanation of social phenomena under study" (Corbin & Strauss, 1990, p. 5). Overall, the GT methodology does not entail strict rules and requirements but provides guidelines with some flexibility (Charmaz, 2014). The specific GT approach applied for this work follows especially the approach as described by Charmaz (e.g., Charmaz, 2006, 2014; Charmaz & Thornberg, 2021).

Table 8: Overview of key qualitative methodologies and their main goals.
Source: Based on Creswell (2008).

Narrative research	**Phenomeno-logical research**	**Ethnographic research**	**Case studies**	**Grounded Theory**
Creating a combined narrative or story from one or more participants' and the researcher's lives	Understanding the essence of a specific lived experience	Observing and describing a specific cultural group	Describing few selected cases in great depth, typically with varied data and over time	Deriving a theory based on the views of participants

GT was chosen because it provides a structured and systematic approach to data collection and analysis. It allows to develop theory grounded in the actual experiences and perspective of consumers rather than just creating high-level descriptions or forcing data into pre-existing frameworks. Putting consumer perspectives and beliefs into immediate focus is a key aim of this research and is well supported by the GT methodology as it allows to understand and explore the subjective experiences of consumers. At the same time, following a systematic process, GT ensures that findings emerging from subjective perspectives and beliefs are overall robust and well-supported by empirical evidence. When comparing GT with other qualitative methodologies, it becomes apparent that GT is the best suited for the research at hand. Looking at the goals of the methodologies (see Table 8), narrative research is not suited for the current work because the research questions cannot be answered via creating stories. Similarly, phenomenological research focuses on specific experiences. Ethnographic research focuses on specific cultural groups, which is not applicable to the research at hand, and case studies focus on specific cases rather than providing more generalizable insights into concepts as is needed for answering the research questions.

Further, GT is specifically recommended "for generating depth of understanding when little is known about a phenomenon" (Flint et al., 2002, p. 103), which applies to this research. For example, in the process of identifying PEF dimensions (if PEF is perceived as multidimensional construct) little is known to start with, but via constant comparison and iterative analysis, GT can help to create increasing depth in identifying and explaining relevant dimensions. Similarly, completely new factors influencing relations between PEF and PQ may get identified and will need to be conceptualized and refined until they are thoroughly explained.

Additionally, GT allows specifically to gather and analyze relations between constructs in detail. This is important for the research questions of this work, as correlations between PEF and PQ are the central focus. This research also aims to identify factors influencing correlations between PEF and PQ. For these influencing factors, it is necessary to understand if they are moderating or mediating the relations between PEF and PQ. With GT, new insights can be explored as they emerge and can be assessed further in the following data collection to ensure overall theoretical saturation in the end.

Finally, GT is a well proven method in marketing research and many authors use GT in research published in leading marketing journals and / or research with a similar focus on PEF

(e.g., Challagalla et al., 2014; Chase & Murtha, 2019; Flint et al., 2002; Homburg et al., 2020; Nenonen et al., 2019; J. C. Nunes et al., 2021; Warren et al., 2019; Wiederhold & Martinez, 2018; Zeithaml et al., 2020). Applying this method follows a recent call in the Journal of Marketing (Zeithaml et al., 2020) for more qualitative, in-depth research that first defines an emerging construct from a participant's viewpoint (e.g., PEF) and then assesses the interrelationships with other constructs (e.g., PQ).

4.2.2. Grounded Theory process

After having outlined the choice specifically for GT as research method, the process applied is explained in more detail in this section. Overall, this thesis applies the core components of GT like rich data gathering, constant comparison, theoretical sampling (stepwise identification of relevant sources based on emerging concepts), first-order / open coding line by line, second-order / focused coding of emerging themes, memoing, development of aggregate dimensions and capturing the dynamic relationships between concepts (Charmaz, 2006; Charmaz & Thornberg, 2021; Gioia et al., 2013). It is important to note that GT is not a linear process but rather a back and forth between data collection and data analysis, and also within the different stages of data analysis (Saldaña, 2013). Typically, data collection and data analysis occur simultaneously, influencing each other (Charmaz, 2006) – which is the case in this thesis. The data collection and analyses follow in every step the quality criteria for qualitative research as proposed in literature (Flint et al., 2002; Zeithaml et al., 2020), namely credibility, transferability, dependability, conformability, distinctiveness, integrity, fit, understanding and generality. In addition, the checklist for quality in GT developed by Charmaz and Thornberg (2021) was considered and helped guide the process and choices made during data collection and analyses. Figure 9 shows a visualization of the steps performed for planning and conducting this research. The numbered boxes (1-6) will be described in more detail in this section as they are the central steps for data analysis.

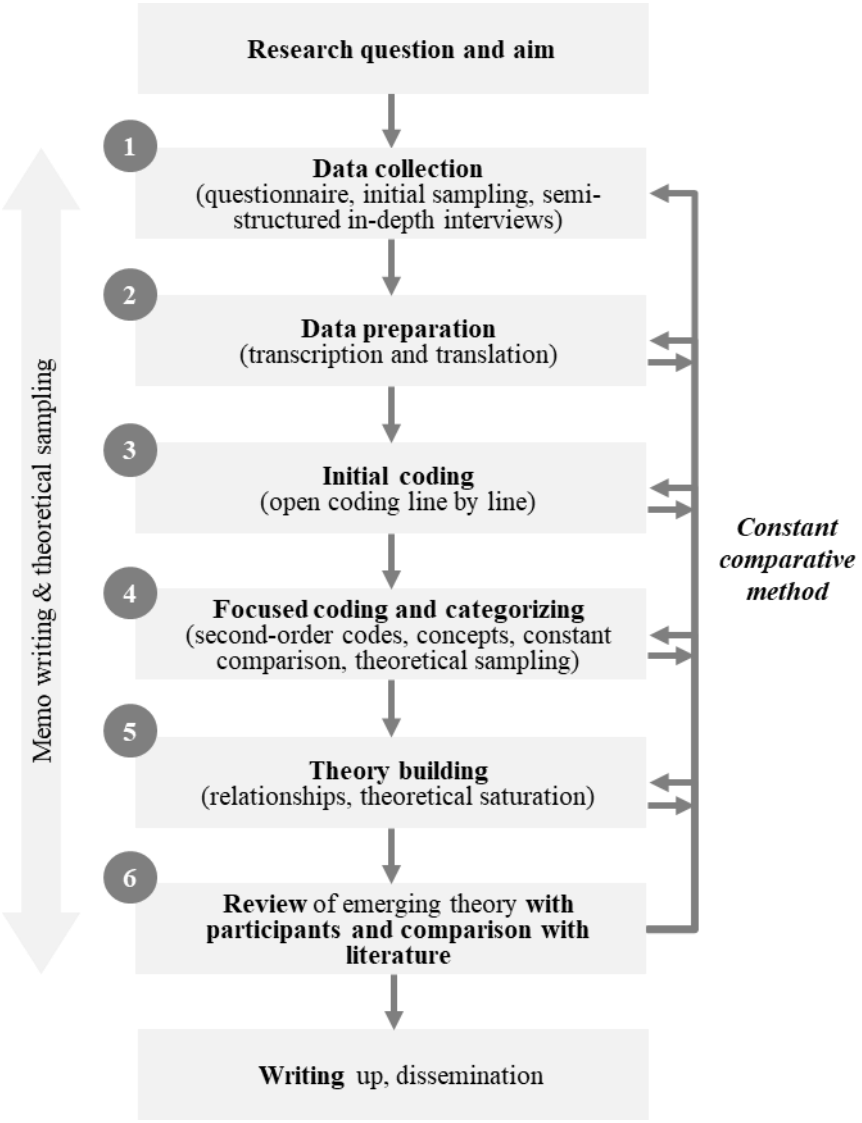


Figure 9: High-level visualization of the GT process as applied in this thesis. Source: Own adaptation based on Charmaz (2014).

4.2.2.1. Data collection

This section explains in more detail the data collection as first step (after the aim and research question) within the GT process described in the last section. The following sections will then guide through and explain the remaining steps within the process. An intensive interview-based approach has been chosen for data collection to gain thorough insights, follow emerging leads and gather rich data as described by Charmaz (2006). According to Charmaz (2006), intensive interviews enable the interviewer to dig deeper than in normal conversations, e.g., requesting explanations, clarifying meaning, shifting the focus, exploring thoughts, feelings and action, or coming back to earlier points. In general, interviews are favored compared to other data gathering methods within GT and specifically the use of semi-

structured, open-ended in depth-interviews is most common to gather rich insights (Goulding, 2002). The semi-structured format uses several broad and open-ended questions as well as follow-up questions digging deeper into emerging topics, clarifying specific meaning, and probing answers. Thus, a semi-structured format is chosen for the research at hand.

In line with the GT appraoch, an initial interview guideline was developed based on the research questions and with the target sample in mind. The guideline intends to help the interviewer during data collection, e.g., to avoid asking suggestive questions (Charmaz, 2006), and to use wording that the participants can easily understand. To develop the semi-structured guide for this research, four other researchers (two professors in marketing with both methodological and subject experience and two PhD holders that had previously conducted interview-based qualitative research themselves) reviewed and gave feedback to a preliminary version of the interview guide. The researchers were identified from personal contacts of the researcher and selected because of their content knowledge of the topic of PEF / PQ or because of their methodological knowledge of interview-based research and / or GT. Other researchers were involved to reduce subjectivity and increase the validity of the data collection – ensuring that the data are relevant and suitable to answer the research questions. The expertise of the researchers also helped to rephrase potentially unclear or leading questions. Based on their feedback, few questions were eliminated while few others were added or slightly adjusted.

In addition, five interviews were conducted using a preliminary version of the guide to further test and refine the interview questions based on participants' understanding and ability to answer the questions. These pre-test participants were identified from the close personal network of the researcher (i.e., friends and family). Interviewees were instructed to voice out their thoughts and highlight if they perceived anything to be unclear or misleading within the interview (e.g., unclear wording, leading questions, logical breaks, etc.). Based on the responses, the initial questions were slightly adjusted in wording to ensure better understanding.

The interview guide is structured into five areas to guide the flow of the interview and ensure that all necessary information and concepts in line with the research questions are explored: 1) Participant characteristics and experience; 2) Comprehension of PEF; 3) Comprehension of PQ; 4) Correlations between PEF and PQ; 5) Wrap-up of the interview. The logic of these areas has been tested in the first preliminary interviews. The first area intends to understand each interviewee and their specific situation, and to engage the interviewee with relatively simple questions at the beginning of the interview. Gathering information about the

participants also helps to put answers into context and to conduct theoretical sampling (by identifying and relating interviewee characteristics). Areas 2), 3) and 4) were included based on the research questions and the central concepts of PEF and PQ. The order of areas 2) and 3) was changed from time to time to avoid systematic influence (e.g., mentioning PEF as PQ criteria because PEF was just discussed). Table 9 shows exemplary questions in the semi-structured interview guideline.

Table 9: Exemplary questions from the semi-structured interview guideline.
Note: Probing / follow-up questions and specifically tailored questions about emerging concepts not included here.

Topic	**Exemplary question**
Characteristics of participant	• Could you please tell me a bit about yourself / your role / your company? • How much experience do you have personally with eco-friendly products?
Comprehension of PEF / PEFPs	• What does eco-friendliness mean to you? • How would you define an EFP? • What makes a product eco-friendly? • Under which conditions can a product not be EF? • Please imagine now that you look at a product. What do you do to assess if it is EF? (What is the process? What do you do next? What are the criteria? What do you pay attention to? Is that everything? What happens if you are unsure about the EF? Which role does […] play?) • Under which conditions do you trust in a product's EF? • How has your perception of EFPs changed over time? • Which pros and cons of EFPs have you observed? • If you compare EFPs and not-EFPS, how do they differ?
Comprehension of PQ	• What is (perceived) product quality for you? • What do you do to assess if a product has high quality? (What do you look for? What makes you choose these factors? What happens if you are unsure about the PQ?) • Which (other) factors do you consider when judging product quality? • Which role does XYZ play for your quality assessment? • If you compare high quality products and low quality products, how do they differ?
Link between EF and PQ	• For you, how (if at all) are EF and PQ of a product related? • How strong is this relation? Why is that? • How have these relations changed over time?

Topic	Exemplary question
	• Do you think there are any other connections between product EF and PQ? If yes, which? When are they positively / negatively / not related? How is this dependent on product characteristics? • When does EF not influence your quality perception? • Can you please give me some examples of high-quality EF products and low-quality EF products? Can you please give me some examples of high-quality non-EF products and low-quality non-EF products?
Wrap-up	• Is there something you would like to add to what we have discussed? • Do you have any other comments or any questions for me? • Can you recommend someone (…) that I should also interview? • Can I get back to you in case of questions that arise during analysis?

In addition to the questions presented in the table, follow-up and probing questions were asked throughout all interviews, for example, "How did you come up with that?", "Why is that?", "What does … mean to you?", "Are there any other …?". Over the course of the data collection, the main questions of the interview guide remained mostly the same, while the follow-up questions changed in line with the emerging framework to collect rich data on emerging ideas and concepts. In later interviews, more focused questions about specific emerging concepts were used, for example about positive emotions:

"*With conscience, we just moved into the direction of emotions, and you mentioned your mood before. Does this also play in important role regarding the topic of EFPs for you?"* (Interview 16); or *"Would you also say that, or how would you say do your emotions or your feeling influence your PQ?*" (Interview 15). Another example is about social friendliness as dimension: "*Does for you, I intentionally called it EF, does a social aspect like fair trade also belong to EF or would you see this separately?*" (Interview 15). During the later interviews, the researcher had a list of emerging concepts in sight to be able to ask about specific concepts if the participant would not bring them up proactively.

The recruitment and initial sampling of participants was planned based on considerations about who can provide the required insights to address the selected research questions. The initial participants were intended to be general consumers, with differing levels of exposure and knowledge about eco-friendliness. Recruitment started via convenience and purposive sampling from the private network of the researcher (e.g., contacts of friends). See

section 4.2.2.5 for more information on the application of theoretical sampling at later stages of the research and section 4.2.3 for more information on the final sample.

4.2.2.2. Data preparation

Data preparation started almost in parallel with data collection by transcribing each interview individually as they occurred. The researcher personally and manually transcribed and where necessary translated the interviews to get immersed in the data. DeepL Translator (DeepL SE, n.d.) was selectively used to support translation but all suggestions were reviewed in detail and adjusted by the researcher as she saw fit. Initial memos were written during these steps to capture early thoughts or observations. The author transcribed all recordings true verbatim, i.e., word-for-word (Zhou et al., 2013). The transcripts created contain each spoken word (e.g., including repeated words when the interviewees were thinking about what to say next). However, "ahm", "uhm", pauses, etc. were not specifically added to the transcripts because they are not important for data analysis in this research context. No interviewee spoke in a dialect that required adjustments or clarification of specific expressions. The next paragraphs describe the further data analysis in more detail and show examples from the coding process.

4.2.2.3. Initial coding

Within GT, once the first data has been collected and prepared, initial coding is started. A *code* "is most often a word or short phrase that symbolically assigns a summative, salient, essence-capturing, and/or evocative attribute for a portion of language-based or visual data" (Saldaña, 2013, p. 3). Initial coding is commonly done via line-by-line coding, i.e., assigning one code per line of transcripts, in a style using present participle to describe the action that is happening (e.g., "trusting friends"). Line by line coding ensures that the researcher is taking a close look at the data, sees different nuances, and can get inspired for new ideas (Charmaz, 2006). In this thesis, after having conducted the first interview and having prepared the transcript, line-by-line coding started within MAXQDA (software for qualitative data analysis). Initial open codes were very closely connected to the interview data and mainly used the participants' language and expressions (called *in-vivo codes*).

During the coding process, additional thoughts, observations, and questions were put constantly into memos where it seemed relevant. *Memos* are "preliminary analytic notes (…) about our codes and comparisons and any other ideas about our data that occur to us" (Charmaz,

2006, p. 3). In other words, memos are "ideas which have been noted during the data collection process, and which help to reorient the researcher at a later date" (Goulding, 2002, p. 65). By writing memos, the researcher can capture emerging ideas, start identifying correlations and move closer towards the developing theory. Table 10 shows an example of initial codes, second order (focused) codes (which will be explained in the next section) and memos written about specific coded incidents.

Table 10: Exemplary interview quotes.
Notes: In vivo codes highlighted in bold. Second-order codes were added during the next stage of analysis and are shown here to highlight the relations between the codes and memos.

Quote Example (initial in-vivo code in bold)	**Second-order codes**	**Memo**
*"I do not really. I probably **trust the label**, and if it says something that indicates that it might be more eco-friendly, I just **feel good about it**. A little better."(Interview 1)*	1. Recognizing PEFP mainly by labels 2. Feeling better / good emotions due to PEFPs	n/a
*"It really **depends on my mood**. But I think in situations where you have two products next to each other and you can clearly see that one is **labelled in a different way** and then you start comparing."(Interview 1)*	1. Depending on time available, awareness and mood 2. Paying more attention to EF when PEFP and CP are directly next	Second time she talks about her mood → linked to how much time / stress but maybe also to emotions? – which roles play emotions? – which kind of emotions are relevant? – for whom are emotions relevant? – situational or overarching?
***"**Positively. So, I think if they **put a lot of thought** into EF, they put a lot of thought into the product in general, so the **quality might be better.**"(Interview 1)*	1. Believing PEFP to be more thought through 2. Positive effect of PEF on PQ	Positive relationship as "they put a lot of thought into the product in general": – positive halo effect? – is it for all products? – when does this not apply? – "they" relates to whom, manufacturers? – is this a common sentiment?

4.2.2.4. Focused coding and categorizing

After initial coding and memo creation about the first data, focused coding starts according to the GT process. Focused coding means that the attention of the researcher is put more and more to codes that "assume greater analytic power" (Charmaz & Belgrave, 2015, p. 3). It allows "to separate, sort, and synthesize large amounts of data" (Charmaz, 2006, p. 11). In doing so, the assigned codes are grouped together and labelled with tentative names for the higher-order codes – which is all feeding back into earlier stages again for constant comparison.

During the data analysis, second-order codes were created first based on the initial open codes and then summarized into tentative categories, before finally being organized into overarching themes. After open coding of five interviews, the author started to combine codes with the same meaning and to create clusters of related content which formed the second-order codes. From then on, existing second-order codes were applied where they fit, and new codes added via open coding where no existing second-order code seemed suitable. Parallel to the coding process, observations, correlations, and ideas to explore were continuously captured in memos. After coding several interviews, the author went back to earlier transcripts and their codes to review them, identify potential updates (e.g., creating new second-order codes by grouping open codes) and look for additional insights based on the increased knowledge over time. This method is called constant comparison and is a key part of the GT process. According to Glaser and Strauss (1967), this means that each coded incident is compared with the previous incidents of the same category across groups or individuals, leading to continuous refinement and definition of the categories. Throughout these steps, further interviewing and transcribing continued in parallel with coding the available transcript(s).

At this stage, the author applied *theoretical sampling*, to ensure information gain was maximized. "Theoretical sampling means seeking pertinent data to develop your emerging theory. The main purpose of theoretical sampling is to elaborate and refine the categories constituting your theory. You conduct theoretical sampling by sampling to develop the properties of your category(ies) until no new properties emerge" (Charmaz, 2006, p. 98). Specifically, two strategies were used to apply theoretical sampling: 1) Identification of participants based on the emerging theory; 2) Identification of questions to ask during the interviews and of relevant data incidences in the conducted interviews. Both strategies are briefly explained in the following.

Regarding 1) the identification of participants, specific participants were approached based on the emerging concepts. Once second-order codes and concepts / categories started to emerge, participant selection was informed by these emerging elements of the theory. For doing that, information already known before conducting a specific interview was leveraged. This was possible during theoretical sampling, because participants were recruited via the researcher's private and social network and asking if her contacts could refer interview candidates that fit specific criteria. In addition, the researcher asked participants at the end of the interview if they could refer further candidates with specific criteria. To maximize information gain, both similarity and divergence regarding those criteria was tried to achieve within the sample. Such criteria included e.g., the level of knowledge about (specific) PEFPs (relating e.g., to emerging concepts of perceived product innovativeness), the general attitude towards animals (relating e.g. to the emerging concept of animal friendliness), or the family status (relating e.g., to the emerging dimensions of PEF or the emerging concept of anticipated warm glow). More detailed information about (theoretical) sampling and the final sample is described in section 4.2.3.

Regarding 2) the identification of questions to ask, the author checked the guiding questions in the interview guide regularly based on the analyses to reflect the developing model and probe tentative categories further. For example, after the emerging category of positive emotions (earlier naming of *anticipated warm glow*) was identified, later interview participants were asked directly about the role of their emotions (e.g., interview 14 and interview 16). In addition, due to the constant comparison between different data and revising earlier interview transcripts multiple times, different questions were asked of the collected data in each revision of earlier interviews. This led to additional insights and coded incidents (i.e., interview quotes) even in earlier transcripts.

Figure 10 shows an illustrative example of the bottom-up coding and specifically of grouping interview quotes / incidents to second-order codes and those second-order codes to categories. Some of the categories were then further grouped into themes. Note that only a very small example of the overall codebook is visualized in the example below for explanatory purposes.

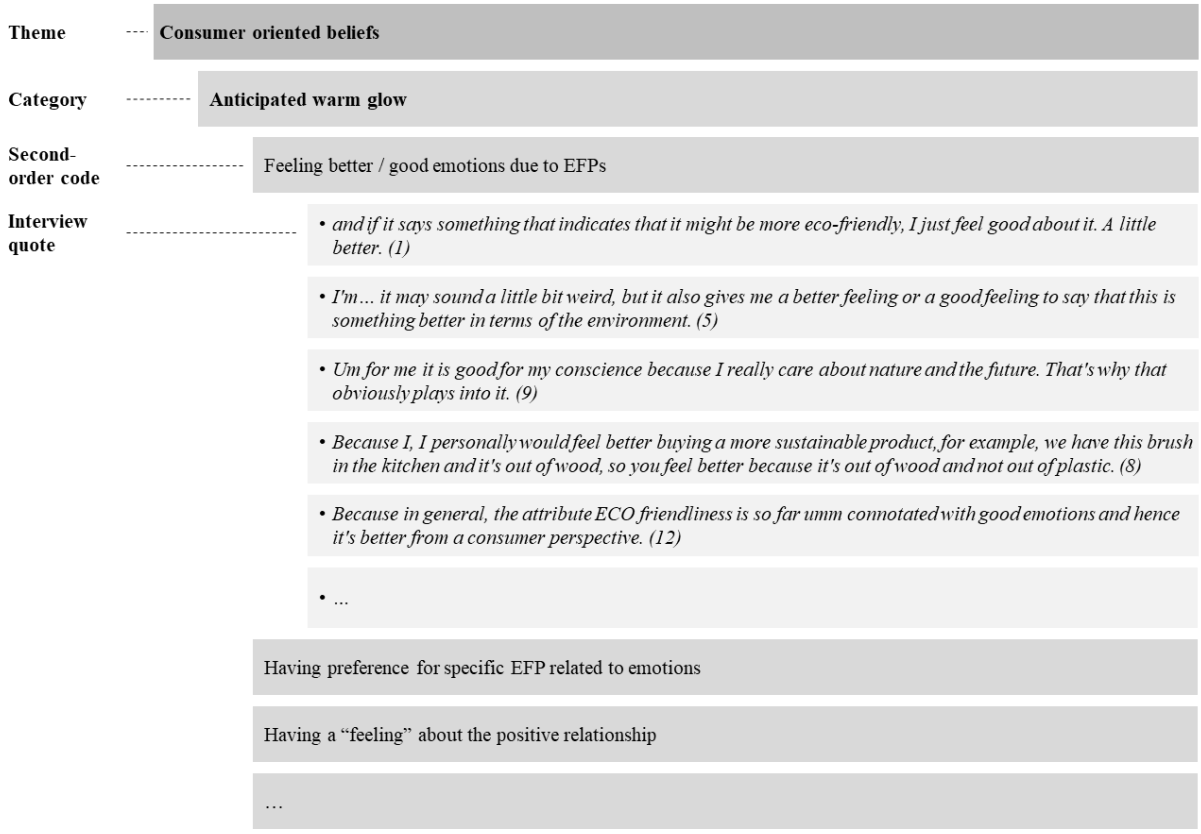


Figure 10: Example of bottom-up coding and analysis.

4.2.2.5. Theory building

Once focused coding and tentative categories have progressed, the correlations between them start to be captured (based on e.g., the context within the interviews, earler codes or memos) to build the theory. The whole iterative process is usually stopped when theoretical saturation is reached, meaning that adding new data does not add new properties to the emerging theory anymore (Glaser, 2001, as cited in Charmaz, 2006). When this stage is reached, the findings are put together and explained based on the codes and memos that had been created. During this process of *theoretical sorting* (Charmaz, 2006, p. 115) the researcher brings together the information collected via the codes, categories and memos developed throughout the research.

Specifically in this work, the researcher started with the concepts of PEF and PQ and then began to diagram and arrange all other concepts around these two core categories. The researcher went back to the research questions and started to build the model bottom-up from the product attributes until the quality perception. All emerging concepts were written onto post-it notes and put on a whiteboard so that they could be easily rearranged, and notes or questions could be added. Whenever the model did not seem to progress, the researcher went

back to the memos, codes, or transcripts to get a new view on the underlying data. In this process, the concept of PEF and how it arises or is identified by consumers was put into focus first. Once the model around PEF was set-up, the focus shifted to visualizing and explaining the relations between PEF and PQ with all the identified influencing factors. The revised version of the model is visualized in section 4.3 where the results are explained in more detail.

4.2.2.6. Review with participants, researchers, and literature

To ensure the quality and fit of the presented emerging theoretical model, the prelimianry model was shared and discussed with five participants who had voiced to be interested in reviewing and discussing results. Participants were shown a visualization of the model, and the researcher briefly explained the included concepts verbally (via phone or in person). Participants then were asked to provide feedback via various open questions, e.g., "How well does the model represent your view on the presented constructs", "Which (if any) parts seem unclear or illogical to you?", "Do you have any other feedback regarding the model or this research overall?". Based on the feedback from participants, the concept of anticipated warm glow was changed from a moderator to a mediator between PEF and PQ. In addition, slight adjustments (e.g., wording updates) were made based on participants' feedback. Beyond participants' feedback, the model was discussed with two researchers, who recently had finished their PhD studies in different academic fields. This step was taken to test the emerging theory against the view and understanding of people who were not involved in providing data input for the model but are overall familiar with research and theoretical models (including e.g., moderator and mediator variables). Their questions and feedback led to more thorough explanations of some concepts to facilitate understanding but did not change the model.

In addition to receiving feedback from individuals, existing research was used to challenge and enrich the emerging findings at this point. Keyword search in EBSCOhost and Google Scholar was used to identify previous research about the emerging concepts – in addition to revising the literature identified in the previous systematic review – specifically regarding the constructs emerging from the qualitative work. This step was taken to understand if there might be any missing aspects and to be able to relate current findings from the interviews with existing insights (e.g., where they are in line with each other versus where and how they are contrasting). For example, the dimensions of PEF identified from the interviews were also tested with elements of PEF found in academic literature (e.g., Boz et al., 2020; Dangelico & Pontrandolfo, 2010; Henninger et al., 2016; Jones et al., 2016; A. T. Nguyen et al., 2020; Van

Bussel et al., 2022). When comparing the identified dimensions with literature, no gaps in the suggested dimensions could be detected, meaning that the dimensions seem to be exhaustive. At this point, the GT process has reached its last step (see Figure 9) and the findings are written up which concludes the conceptual work. In the next section, the final sample of the qualitative research is described before the results are presented in section 4.3.

4.2.3. Final sample

The final sample is a theoretical sample of 22 participants. Interviewing was stopped after four consecutive interviews added no new information to the model (theoretical saturation), resulting in the sample size of 22 interviews. Beyond saturation, the sample size of this work lies within the typical range for GT research: Thomson (2010) investigated sample sizes of interview-based GT research and out of 100 identified studies, 44 used less than 20 participants, and an additional 33 studies used between 20 and 30 participants (overall average sample size of 25 participants).

Initial recruitment took place using convenience and purposive sampling (using also snowballing), in line with the GT process as described by Charmaz (2006). Later participants were chosen based on their expected ability to explore emerging concepts and to add new insights (theoretical sampling). Recruiting of participants in line with theoretical sampling took place via personal contacts, references (e.g., from previous interviewees) and social media.

The dimensions / criteria used for theoretical sampling were chosen based on findings in previous literature as well as emerging ideas throughout the interviews. The potential relevance of age, and gender was derived from previous research (Abbey et al., 2015; Tong & Su, 2018; Van Doorn & Verhoef, 2011) as outlined in section 3.4.2.2. Occupation was used as a proxy for both income and assumed knowledge about EF (in professional context). In addition to occupation, discussions with (potential) participants took place to identify their knowledge about EF (which has been shown as relevant by e.g., Abdulla et al. (2024), Farías (2020) and Y. Wang & Hazen (2016)). Similarly, personal importance of PEF, concern for the environment and attitude towards PEFPs were factors considered in participant selection and were mostly derived from discussion with (potential) participants.

Similarly, more specific considerations about individual emerging concepts were used at different points during participant selection. Four examples of those considerations and applications of theoretical samples are outlined in the following: As first example, attitude towards animals was considered for exploring the emerging concept of animal friendliness, and

while participant 5 was not particularly fond of pets, participant 6 was selected partially for her professional ties to animals. As second example, participant 9 was selected because she lives in a rather rural / remote area, thus has much less access to PEFPs than previous participants and because she recently became a mother which might have shifted some of her perceptions. Both factors were considered relevant to explore several emerging concepts, e.g., positive emotions or PEF dimensions like climate friendliness or naturalness. As third example, participants 12, 13, 18 and 20 were selected due to their professional background in specific industries like grocery retail, personal care, clothing, and raw material production (amongst other factors like differing international exposure). Having varying degrees of knowledge about specific product types was considered relevant to substantiate the emerging theory as the intent of this work was to develop a theory applicable cross-industry (within consumer goods). In addition, it helped to explore emerging categories like product innovativeness from people with different experiences and knowledge. Beyond a focus on specific emerging categories, participants were selected to be either rather different or similar to each other (as much as it was possible to understand a priori). As fourth example, participants 11 and 14 were different in terms of their focus and relevance of PEF in their everyday lives as well as in their lifestyles, while participants 14 and 15 were very similar in those elements.

Participants were asked about language preferences and decided between English or German before the recoding was started. All interviews conducted in German were transcribed in German and then carefully translated to English. 14 participants (64%) chose Englisch as interview language and eight (36%) participants chose German. Language of the interview is included in the overview below to show which interviews and thus also quotes presented in the following result section were translated by the author. All participants either have a European nationality (mainly German) or have been living in Germany for several years.

The final sample consists of 17 adult general consumers (77%) as well as five practitioners (23%) in the sustainability, marketing, and brand management areas of consumer goods companies and retailers or founders. As only adults were considered for this study, participants' ages span from 18 years to mid-60s. Specifically, nine participants (41%) are within the age group of 21-30 years, five participants (23%) within the age group of 31-40 years, three participants (14%) each within the age groups of 18-20 years and of 51-60 years and the remaining two participants within the age groups of 41-50 years and of 61-70 years, respectively. Both sexes were represented across age groups and occupational areas, in total 13 women (59%) and nine men (41%). Participants have different occupational backgrounds, e.g.,

student, marketing / brand manager, entrepreneur, sports instructor, secretary, management consultant or unemployed. Table 11 shows selected information about each interviewee aiming to support general understanding of the sample.

Table 11: Descriptive details about interview participants.

ID	Language	Age group	Sex	Category	Occupation
1	English	21-30	Female	Consumer	Marketing expert (professional services)
2	English	21-30	Female	Consumer	Pharmaceutical researcher
3	English	51-60	Male	Consumer	Project manager (technology)
4	English	21-30	Male	Consumer	Entrepreneur (consumer goods)
5	English	31-40	Female	Consumer	Investor relations (consumer goods)
6	English	41-50	Female	Consumer	Horse riding instructor
7	English	21-30	Male	Practitioner	Entrepreneur (consumer goods)
8	German	21-30	Female	Consumer	Marketing expert (media)
9	German	31-40	Female	Consumer	Secretary (healthcare)
10	English	21-30	Female	Consumer	Student (management)
11	English	21-30	Female	Practitioner	Sustainability consultant
12	English	31-40	Male	Practitioner	Strategy consultant (consumer goods)
13	German	21-30	Female	Practitioner	Marketing expert (consumer goods)
14	German	18-20	Male	Consumer	Student (vocational training)
15	German	18-20	Female	Consumer	Student (philosophy)
16	German	21-30	Female	Consumer	Project manager (technology)
17	German	61-70	Female	Consumer	Dentist
18	English	31-40	Male	Practitioner	Marketing expert (consumer goods)
19	English	31-40	Male	Consumer	Product designer
20	English	51-60	Male	Consumer	Entrepreneur (industrial goods)
21	English	18-20	Female	Consumer	Gap year / skiing instructor
22	German	51-60	Male	Consumer	Unemployed

Interviews were conducted in a one-on-one setting. Often, researchers conduct GT interviews face to face (Goulding, 2002), however for broader geographical coverage and due to the ongoing global pandemic (caused by the coronavirus disease 2019) during data collection, most interviews were conducted via video conferencing (20) and only two interviews were conducted in person. All interviewees were informed about the procedure and nature of the study. Everyone was asked for voluntary consent and agreed to have the interview recorded. During the interview, participants were encouraged to share their video to enable non-verbal observations during the interviews but for privacy reasons videos were not saved even if the camera was enabled (only audio recordings were saved). All interviewees were free to choose where to conduct the interview from and chose a private and quiet surrounding for the interview, mostly in their own home. For the two interviews conducted face to face, the researcher met

the respective interviewees in their own home. Interviews lasted on average 55 minutes with a minimum duration of 36 minutes and maximum duration of 95 minutes. The final dataset consists of 275 pages of transcripts with more than 155,000 words (English transcripts).

4.3. Qualitative research results

After the detailed description of the qualitative method and final sample, the results from the GT analysis are presented in this section. It follows the order of the research questions that this study is investigating, with section 4.3.1 addressing definition and conceptualization of PEF, section 4.3.2 explaining the results in terms of dimensionality of PEF and section 4.3.3 describing the constructs identified as influencing the PEF – PQ relationships.

4.3.1. Defining and conceptualizing PEF from consumer perspective

This section starts with the development of a PEFP definition and then describes the product attributes and other influencing factors consumers consider in assessing PEF. Interview quotes are included in italics for illustrative purposes.

4.3.1.1. PEFP definition

As input for creating a consumer definition of PEF, participants were asked to define PEF / PEFPs in their own words. In the literature, the term EF is often used synonymously with related terms like environmentally friendly, ecological, green, sustainable or ethical (e.g., De Medeiros et al., 2014; Gleim et al., 2013; Gleim & Lawson, 2014; Lin & Chang, 2012; Liobikienė & Bernatonienė, 2017; Sdrolia & Zarotiadis, 2019; W. Wang et al., 2017). This was also found in the interviews, where participants used several words like EF, sustainable, environmentally friendly, responsible, or green. For example, one participant chose several different terms throughout the interview:

"Okay, so I know from my partner that they have, they, they're very responsible, environmentally responsible. (...) I don't perceive as environmentally friendly at all. (…) *I think that that is contributing a little bit to becoming more green or more eco-friendly. (...) I would say, because you're, you're not just promoting for sustainable ways of manufacturing and managing waste (...)."* (Interview 19)

Frequencies of word use cannot be assessed in this research, as the wording choice of participants who opted for a German interview cannot be directly compared with the originally English interviews due to potential translation fuzziness or differences.

To derive a PEFP definition from consumer perspective that is still related to the academic perspective, the working definition introduced in chapter 2 serves as basis and was adjusted to reflect specific terms and content that participants mentioned. The resulting definition is:

A PEFP is a product that consumers perceive to have a more positive impact or a lower, not long-lasting, or preferably no harmful impact on the environment (non-living and living things, including animals and humans) across its whole life-cycle (before, during and after use) compared to alternatives that fulfill the same consumer need.

The content of the definition can be broken down into several elements, specifically valuation, scope, and reference (in the order they are mentioned in the definition). Table 12 summarizes the key elements and the following paragraphs further describe them.

Table 12: Key elements of the PEFP definition.

Element	**Representation in definition**
Valuation	→ positive or negative (harmful) impact
Scope	• How much / how long lasts the impact: → more for positive; → lower, not long lasting, or not at all for negative • Who is affected: → non-living and living things, including plants, animals and humans • When does the impact occur: → across whole life cycle
Reference / baseline	→ compared to alternatives that fulfill the same consumer need

The participants described positive and negative impact on the environment as relevant. They mostly used the term *harmful* and agreed that no such impact should occur. However, participants understood this is difficult and included that the negative impact of a certain product can be lower than the negative impact of relevant alternatives, or the negative impact can be only temporary and not long lasting for a product to still be considered PEF. Most participants explicitly included animals and humans in their definition, in addition to plants and non-living things (e.g., water). Thus, PEFPs cover ecological and social aspects. Participants also considered the whole product life cycle, either by directly using that term or by describing its different phases (before, during and after use). Lastly, participants compared products to alternatives with the same purpose, fulfilling the same need, e.g., different hair shampoos and hair soap (purpose: To clean and care). These alternatives can be earlier versions of products,

meaning that PEF for the same product can change over time. For example, if other products fulfill higher PEF standards while the product in question has not changed then it is considered less PEF than it used to be. As one participant put it: *"The standards have also risen. And... if somehow things are still not at the level of five years ago or so, then you think to yourself, yes... You are not at all environmentally friendly, and also still behind"* (Interview 14). Additionally, "*It's also nothing, I would say, written in stone, and then it lasts forever. It will also change throughout the years"* (Interview 3). Accordingly, participants saw PEF as something relative on a scale from not PEF to fully PEF instead of a yes or no decision. They also highlighted, that regardless of the product, reducing consumption is key to more EF because *"It would of course always be more EF not to grow the coffee and not to consume it (...) because everything that we buy and consume harms in some way, um the resources and the earth"* (Interview 16).

4.3.1.2. Product attributes

Product attributes play a crucial role for consumer perceptions (Bangsa & Schlegelmilch, 2020; Steenkamp, 1990; X. (Shane) Wang et al., 2022; Zeithaml, 1988). Product attributes here mean the "relatively concrete, objective properties" of a product, that can be observed by consumers (Johnson, 1989, p. 598). This is in line with what Zeithaml (1988) calls *lower level attributes*, for example the brand name or the price of a product.

For PEFPs, participants reported to consider multiple attributes at the same time. This is in line with earlier research, for example that consumers are willing to pay more for products having multiple EF attributes (Zander & Hamm, 2010). Only on few occasions, participants reported to base their judgement on a single attribute, e.g., when being highly familiar with a brand. Participants said that some specific attributes, like containing palm oil, rare-earth metals or even plastic, can prevent EF perceptions regardless of other attributes. For example, one participant said: *"I read that palm oil… firstly, it is not good und secondly it is supposed um, there are some… they clear some forests for those palms to be put there which then…. Yes, then the palm oil… where they can extract it. And that is why there arise long term damages to the environment, also regarding the climate and so on. There is a kind of monoculture which they create there by putting the palms everywhere there. I don't use it anymore, no. You find it a lot in sausages and so on. Um, that is, no it is out of question."* (Interview 17)

In general, the more coherent the impression across attributes, the more trust is given, which helps preventing greenwashing perceptions: *"So, if it says it is EF, but it is packaged in plastic and the label looks a little sketchy and I am not sure, then I might not really trust it. But*

if it has this label on it and it is sold at for example a "Bioladen" [organic store], and the packaging looks really EF at the first glance, then I might trust it more." (Interview 1)

Previous research also found that the centrality of an EF attribute for the product as a whole impacts PEF (Gershoff & Frels, 2015). *Centrality* is "the degree to which an attribute is integral" (Gershoff & Frels, 2015, p. 97) in defining a specific product, e.g., food ingredients are typically much more central than food packaging. In the following paragraphs, the attributes that participants described as relevant for PEF are described one by one.

Origin. Most participants said their EF perceptions are heavily influenced by where a product has been made geographically. Product origin is recognized as important attribute in current EF research (e.g., Van Bussel et al., 2022). The origin is considered important because it provides for example information about potential emissions due to transport and about potential production conditions that are associated with specific locations:

"Then you see where it is produced and umm how, because then the logistics comes from me, always into my mind. Yeah, do I have to fly this product from all over the world, is it produced… um ethical, like people or children work or whatever comes to your mind." (Interview 20)

"So, if it says somewhere that it is produced in Poland or somewhere like that, how is it there with the standards for the environment, probably rather difficult." (Interview 14)

The closer the production of a product is to where consumers see or purchase it, the more EF they perceive it to be. To simplify assessment, consumers can think in different stages, namely local / regional, national / German, EU-made or from somewhere else.

"Umm, I mean the more local it is, in my head the more EF it is." (Interview 19)

"So, then I want that it is also written where it is from. So, is it a German product, an EU product, or does it come from anywhere else?" (Interview 9)

Ingredients / materials. Ingredients and materials are considered important for PEF. Contributing to multiple abstract dimensions, participants believed certain materials or ingredients to be more EF, e.g., more natural, and chemical-free, using minimal virgin resources (e.g., rather recycled content), not containing animal substances (e.g., vegan) and easily recyclable. The importance of ingredients for EF perceptions has also been shown in previous research (Magnier et al., 2019). Especially plastic content in a product is detrimental for PEF, while organic materials are seen as EF by most participants. Less ingredients in a product can also contribute to PEF and some participants mentioned the sourcing of raw materials.

"EF is yes, it is based on, on, on sustainable material, it means like recycled material, it means organic material or natural material which, which grows back. And stuff like that that's EF, where the environment is not harmed by, by, you don't apply chemicals which you, yeah which you actually don't need." (Interview 7)

"But I think, overall, one can say that the less ingredients are in the product, the more sustainable they could be." (Interview 10)

"There are materials, where you have some things to… how they get developed or how they get mined or how they get dug from the earth or something…" (Interview 17)

Brand image and size. The product brand plays a big role for most consumers, for some even the biggest role, at least if they know a brand already. Consumers expect the whole brand with its purpose, mindset, and all business efforts to reflect EF. This matches previous findings, e.g., that the perceived motivation of the brand to act EF impacts PEF (Gershoff & Frels, 2015).

"I would first go for the brand because I know which brands, Patagonia, Vaude, are a lot, like are relatively EF." (Interview 5)

"It depends on if the whole company has the image of being EF or not." (Interview 20)

"I think I would tend to personally go for brands that have been... or that I have been, I have perceived them as EF from the start and with all their products." (Interview 11)

Regarding size, consumers believe especially in EF of smaller companies and doubt EF claims of big companies. This effect has also been reported in previous literature (Davies & Gutsche, 2016; Wallach & Popovich, 2023). Reasons reported in the interviews are 1) that consumers believe bigger companies want to keep their conventional non-EF business and profits, 2) that it is more difficult for bigger companies to become EF, 3) that *"less eco-friendly production is more massive and simply can supply enough"* (Interview 2), and 4) that ultimately, they are more likely to engage in greenwashing for PR purposes.

"And so, for me, it's actually, I would say the, the less bigger the brand the more probable is true sustainability of the product." Interview

"But I would also claim that smaller startups that just emerge maybe get more trust from me personally, because some have really this sense of being EF and sustainable at their heart what I mentioned before, that they built their whole business with the background of being EF up with a purpose of being good for the environment." (Interview 10)

"I'm always a bit suspicious of those organic supermarket brands that actually produce again only in mass. (...) Um so I think organic cotton is the same joke as organic food, because

it actually gets produced in bulk, which then for me is not really sustainable again." (Interview 9)

Point of sale image. Many participants talked about the store where they see a product and its influence on perceptions. While most examples were about grocery stores, some mentioned online clothing platforms or drug stores, showing that image has an influence offline and online. Stores that offer only PEFPs and with the right physical environment support PEF. The importance of aligned images (store and product) for PEF became apparent.

"Bio supermarkets yeah that, by its nature, they say everything that we have, everything that we offer follows some of these EF criteria. Otherwise, we would not offer it. So that makes it easier." (Interview 3)

„For example, the typical discounter or you know around… store around the corner, high shelves, a lot of things in there, everything packed. I don't feel it's EF. However, other stores, where they, you have more space, everything is wooden everything is maybe in a glass char so you have to bring your own let's say boxes, where you can fill the stuff in... I feel much more EF in this case (…)." (Interview 5)

"So, if a brand decides for example to be listed exclusively at a retailer who promotes sustainability heavily, then it is automatically… so then an alignment takes place right away. And then also the brand is perceived, I would say, faster as being sustainable." (Interview 15)

Packaging. All participants found packaging to be important for PEF. Especially the material, size and design were discussed. Similar to product material, packaging material should be natural, recyclable, etc., which should be reflected in the design (e.g., natural colors). Additionally, consumers found: The less packaging, the better. This is in line with previous research, e.g., with structural and graphical cues identified by Magnier and Crié (2015), with the dimensions described by Nguyen et al. (2020) and with the categories by van Bussel et al. (2022).

"Okay, EF, from my perspective, means umm no or little packaging." (Interview 5)

"I guess the less plastic that is used in, in packaging, the better." (Interview 18)

"There are definitely colors that look like you think it might be more EF, it is like more muted colors compared to like flashy colors that look kind of toxic." (Interview 1)

Eco-Labels. *Eco-labels* are informational cues, e.g., certificates or seals, that are supposed to ensure the consumer about environmental qualities of a product (Atkinson & Rosenthal, 2014), influencing consumer perceptions (e.g., Dangelico & Vocalelli, 2017;

Magnier & Crié, 2015). Given the broad definition of PEF to include e.g., also social and animal friendliness, a variety of labels like organic, fair trade, vegan, carbon free, etc. is considered in this thesis as eco-label. In line with previous research (D'Souza et al., 2006), participants had a hard time understanding and trusting labels, with some people not trusting any labels.

"But I'm not aware of those that are like real certifications, that don't have any loopholes or things like that." (Interview 10)

"It's also one of the largest weaknesses of this market, that has, that it has due to the jungle of certifications and labels that is out there and there's a confusion amongst the consumer really what is now good and what is bad." (Interview 12)

Some consumers focus on specific labels they know more about and where they believe the issuing party to be trustworthy. Being issued by a governmental institution, independent organization, or NGO, together with being transparent about the assessment and using a scientific approach increases trust in labels. The need for differentiation between different sources of labels is in line with previous research (e.g., Pedersen & Neergaard, 2006).

"I actually check who it is, umm what it, what is the organization, the organization that gets that certificate. And if that sounds or looks good to me, then I, then I trust them." (Interview 4)

"For instance, label created by charity that I trust, so for instance for cosmetics again, there's this leaping bunny label that is, that has been created by this NGO called cruelty free international." (Interview 11)

Price. There was a great consensus between participants that PEFPs are rightfully more expensive than conventional products. For some, this is related to CPs being too cheap:

"I don't understand that the other prices are so low. Because there's so much work behind that and then some products are just too cheap for all the material and all the logistics and everything that's behind it, so I think the sustainable products are often like a fair price or good, or the price that it should have." (Interview 21)

Consumers mentioned several reasons for higher prices, e.g., higher initial investments for producers, more effort and higher standards during production, better materials or ingredients, or lower sales volumes. They believe price to be related to most or all the abstract dimensions. Thus, when a product is perceived as too cheap, consumers lose trust and doubt its PEF. Some use price as sense check if the PEF suggested by other attributes seems realistic.

"Well, if you really think the product is more EF, then it does make sense. Because it also costs to be somehow… more EF and then it makes sense." (Interview 14)

Product design. Product design was repeatedly brought up as contributing factor to PEF judgments, especially color was mentioned as a potential first (conscious or unconscious) cue for PEF. Specifically, participants highlighted more muted, earthy colors, like green, blue, or brown or associated with PEF. Beyond color, product size was mentioned as relevant design element. The role of product design has also been shown in previous quantitative research (e.g., Dixon & Mikolon, 2021; Pancer et al., 2017; Petersen & Brockhaus, 2017). Overall, the findings of these authors seem in line with the current research.

"You just take... or green products, or brownish, in some form natural-colored products stand out in the context of sustainable products of course more than the squeaky red or squeaky yellow. So, there you look, I look now personally in search of sustainable products, I would look there first." (Interview 9)

"So, first, I would have the general impression with regards to the packaging, with regards to the colors, to the size of the product and the ingredients." (Interview 10)

"And the design does not, for most of these products does not really matter that much, but maybe subconsciously a little, too." and later *"There are definitely colors that look like you think it might be more EF, it is like more muted colors compared to like flashy colors that look kind of toxic."* (Interview 1)

"(…) packaged in something green or brown and that's basically everything that comes to mind just yep okay." (Interview 2)

"(…) for example, apples, most of the time they are smaller. But that's absolutely acceptable now because they look different if you compare them directly, side by side." (Interview 3)

However, greenwashing seems a concern arising from product design, specifically from the coloring used: *"There are definitely different design, and you can see when something wants to seem EF, I think. So, there is always a lot of green involved and some flowers or plants somehow included and organic and… however all of that is called. But this is not such a criteria for me, because I think, yes these are somehow… everyone can do this if they want to, such a design. And there are no, or often no stamps and really meaningful label or so that confirm it. And simply because something is green, it does not mean that it is EF. You probably have a second look and then see, ok, it's more environmentally friendly... but it is not a real argument for me now".* (Interview 14)

4.3.1.3. Overarching influences

This section describes the factors that participants described to influence their eco-friendliness perceptions and that may change how the attributes described in the previous section impact PEF for a given product, for a given consumer or in a given situation.

Product type / industry describes in this thesis a similar group of products, e.g., food and beverages, clothing, personal care, etc., and smaller sub-groups of these. Which type a product belongs to can impact overall PEF or its dimensions. Specifically, consumers prioritize different dimensions for PEF of different product groups. For example, while naturalness is important for personal care or food products, it is less relevant for electronics. At the same time, repairability is important for electronics but not really for personal care and food products. Thus, participants adjusted their expectations to different product types so that sometimes they struggled to generalize or rather gave concrete examples.

"So, depending on the, on the segment, I would pick my criteria" Interview (4).

"Um it's hard to like make a generalization about like products in general, obviously, because there's just like so many I'm trying to think of like examples…" (Interview 11).

Additionally, some product types are by default being perceived as more or less EF. For example, meat or fast fashion are on the low end of PEF, regardless of their individual attributes:

"It does depend, because something that is for example the meat industry is always going to be not EF in itself so it be... some meat product could be more EF than another one, but it is still at its core not EF. And then some other industries like, I do not know, I mean a clothing brand that only creates clothes from already created materials or whatever could be completely EF or way more than it is possible for other industries" (Interview 1).

Perceiver characteristics. Characteristics of the perceiver influence PEF in multiple ways. First, they can affect how PEF dimensions are formed based on the product attributes and second, they can change how the dimensions are integrated into the overall PEF. The identified perceiver characteristics can be grouped into demographic and psychographic variables. Some important characteristics are described in this section, but not all characteristics that can possibly influence PEF are mentioned, as it was not the purpose of this research. Based on the interviews, two main demographic variables have been identified: Age and having kids. In both cases, the differences were not that any group perceived in general more or less EF but that they were focusing on different attributes and dimensions or had different approaches. For psychographic variables, some of the most important variables were knowledge / experiences

with PEFPs, concern for environment, and general product preferences. For example, people who have more knowledge about certain attributes or dimensions due to their education and work, rather focus on these for EF assessment than on others. Similarly, knowing more about the drawbacks of labels or certifications makes consumers less likely to consider them: When asked how to recognize a PEFP, one participant answered:

"In the past I would have said by the labels (...). So, with the fair, fair trade label we have actually discussed it in school, I think in 7th or 8th grade in geography or also then with those fair fish, there is a documentary on Netflix about fishery and so this label actually says nothing. There is no fair fish trade. I think also through YouTube I actually know so much." (Interview 15)

Further psychographic variables that can impact PEF are motivation, emotions, or trust in specific information sources, e.g.,

"Sometimes you have just seen a documentary about the effects of some of these products and feel more motivated to do something about it and then sometimes you just forget about it, or you are in a rush." (Interview 1)

Situation / circumstances. There are many circumstances or situations that can be relevant for PEF and influence the perception process: Marketing efforts related to a product, the source of EF information, current societal trends regarding EF, presence and opinions of other people and many more can influence PEF:

"So, it's more what I know or hear from people that are a little bit more experienced than me in that part. Or from yeah maybe YouTube or Netflix shows that I watch" (Interview 19). Another participant does research about ingredients *"when I'm really decisive about healthy life and preserving the planet and everything (...)."* (Interview 2)

Many participants recognized that EF has become a trend and big topic rather recently, typically within the past ~5 years. They believe that with EF being a trend, things are changing, including their perceptions. Especially being more confronted with the topic of EF increases awareness and influences perceptions and behaviors.

"It's become more of a mainstream thing as well, so it's not the purely… How would we say in German "Grüne and hippies" stuff anymore, but it's way more arrived in the center of our society." (Interview 5)

"But I think in situations where you have two products next to each other and you can clearly see that one is labelled in a different way and then you start comparing." (Interview 1)

4.3.2. Understanding the dimensionality of PEF

This section describes the results from the interviews in terms of conceptualization and specifically dimensionality of PEF as described by consumers. Findings from the interviews indicate that PEF is not formed directly in the minds of consumers but rather via several dimensions (like naturalness or social friendliness). Thus, PEF can be conceptualized as *multidimensional construct*, which is defined as follows: "A construct is multidimensional when it refers to several distinct but related dimensions treated as a single theoretical concept" (Edwards, 2001, p. 144). Each dimension reflects a unique aspect within the broader construct (Polites et al., 2012). In the following, the dimensions are defined and described one by one before additional considerations about the interplay between specific dimensions are explained.

4.3.2.1. Description of PEF dimensions

In total, eight distinct dimensions of PEF were identified. Table 13 shows the definitions of each dimension with the following paragraphs describing the dimensions in more detail. Note that the definitions are all phrased in a way to reflect consumer perceptions rather than objective measures (e.g., liters of water used, or tons of CO2 emitted). At the end of this section, Table 14 then illustrates the dimensions via exemplary quotes from the interviews. To ensure validity of the identified dimensions, preliminary results were compared to dimensions and their conceptualizations as found in literature (e.g., Boz et al., 2020; Dangelico & Pontrandolfo, 2010; Henninger et al., 2016; Jones et al., 2016; A. T. Nguyen et al., 2020; Van Bussel et al., 2022). No relevant gaps or contradictory evidence could be found.

Table 13: Dimensions of overall PEF.

Dimension	Definition
Perceived naturalness	The extent to which a product and its components are perceived as belonging to the natural environment, not using chemicals or artificial substances, being minimally processed, and being produced with traditional methods.
Perceived social friendliness	The extent to which a product and its components are perceived as ensuring the wellbeing of humans (directly and indirectly involved).
Perceived animal friendliness	The extent to which a product and its components are perceived as ensuring the wellbeing of animals (directly and indirectly involved).
Perceived climate friendliness	The extent to which a product and its components are perceived as not contributing to climate change and global warming (via greenhouse gases).

Dimension	Definition
Perceived resource efficiency	The extent to which a product and its components are perceived as using or destroying minimal resources (like water, energy, raw materials, etc.).
Perceived durability	The extent to which a product and its components are perceived as lasting over a long period of time without defects (product life).
Perceived repairability	The extent to which a product and its components are perceived as effortless to bring back to a desirable state after defects.
Perceived reuse- and recyclability	The extent to which a product and its components are perceived as having the potential to be used again after initial consumption for the same or other purposes rather than creating waste.

Perceived Naturalness [Nat] was the most mentioned dimension to describe PEF. Participants described it with aspects like not containing chemicals or toxins, not being synthetic (including plastic), being organic, being less produced, being biodegradable / compostable, growing back, supporting biodiversity and being in season / fresh. The definition of Nat as the "extent to which a product and its components are perceived as belonging to the natural environment, not using chemicals or artificial substances, being minimally processed, and being produced with traditional methods" was based on the terms used by participants as well as on previous research about Nat (e.g., Jorge et al., 2020; Román et al., 2017). In comparison to previous research on Nat relating to PEF, participants mentioned Nat as relevant for multiple product categories, while published research focuses mainly on Nat in the context of food products (e.g., X. Chen et al., 2018; Kahraman & Kazançoğlu, 2019; Magnier et al., 2016; Michel & Siegrist, 2019; Pula et al., 2014; Román et al., 2017; Siegrist & Sütterlin, 2017; Staub et al., 2020; Steptoe et al., 1995; Stremmel et al., 2024). Additional evidence for the relevance of Nat was provided by Steenis et al. (2018), who found a positive correlation between PEF and Nat of packaging.

Perceived social friendliness [Soc] is considered a key element of PEF by many participants. Several did specifically highlight the importance of this dimension. According to the participants, it mostly concerns the production conditions (e.g., fair wages, no child labor). When talking about Soc, terms used include society, social justice, people, fair trade, humans, social responsibility, etc. The definition as the "extent to which a product and its components are perceived as ensuring the wellbeing of humans (directly and indirectly involved)" was developed based on participants' input and tries to generalize so far that all their points are being considered. The concept of Soc also plays a role in previous PEF-related research (e.g.,

Ammann et al., 2024; Anagnostou et al., 2015; Bürgin & Wilken, 2021; Davies & Gutsche, 2016; Magnier & Crié, 2015; Stremmel et al., 2024).

Perceived animal friendliness [Ani] was mentioned regularly by participants as being part of overall PEF, mostly regarding not being harmful to living creatures, providing proper living conditions for animals (e.g., having enough space, being cared for), not using animal-related products at all (vegan) or not doing animal testing. Participants not only considered land-animals but also living beings in the sea. The definition as the "extent to which a product and its components are perceived as ensuring the wellbeing of animals (directly and indirectly involved)" was developed based on participants' input and in alignment with the definition of Soc. Additionally, this definition is in line with a recent definition that "the animals should feel well, they should function well and they should lead a natural life (express their natural behavior)" (Achabou et al., 2020, p. 662). The concept of Ani (also referred to as animal welfare) is used regularly in context of PEF related research (e.g., Achabou et al., 2020; Akaichi et al., 2019; Ammann et al., 2024; Blokhuis et al., 2003; Chang & Chen, 2022; Janßen & Langen, 2017; Lindeman & Väänänen, 2000; Staffolani et al., 2022; Zander & Hamm, 2010). Like the concept of Nat, Ani tends to be investigated mostly in the context of food products.

Perceived climate friendliness [Cli] relates mainly to emissions caused during production, transport, usage, or after-life of consumer products. Mostly CO2 was mentioned, but also methane, or air pollution in general. The definition as "the extent to which a product and its components are perceived as not contributing to climate change and global warming (via greenhouse gases)" was partially developed based on participants' wording, however their most used term was CO2 which is not used in the definition but rather the broader term greenhouse gases. Cli and especially carbon footprint has been a relevant concept in PEF-related research for many years (e.g., Ammann et al., 2024; Feucht & Zander, 2018; Hartikainen et al., 2014; Janßen & Langen, 2017; Otto et al., 2021; J. Shi et al., 2018; Wong et al., 2020).

Perceived resource efficiency [Res] concerns most participants. They mentioned resources in general, or specifically water, energy, land, or labor use, and the use of recycled / re-used materials instead of "new" or unused raw materials and not producing in mass to save resources. This theme fits participants' beliefs that not producing and consuming any products is better for the environment as it saves resources. Like the other dimensions, the definition as the "extent to which a product and its components are perceived as using or destroying minimal resources (like water, energy, raw materials, etc.)" was chosen based on the examples provided by participants. Res has been also used in previous research (e.g., Blasch et al., 2019; X. Chen

et al., 2018; Geldermann et al., 2016; Kasulaitis et al., 2021; Marzouk & Mahrous, 2020; Sonnenberg et al., 2011; Stremmel et al., 2024; Tukker, 2015; Visser et al., 2018).

Perceived durability [Dur] is for many an important part of PEF. It has been recognized in literature as dimension of PQ (e.g., Hazen et al., 2017). However, participants in this qualitative research specifically linked Dur with PEF in a way that more durable products are more eco-friendly to them. Dur has a different importance for different product types, e.g., it is less relevant for food than for example clothing items. Based on the input from participants, Dur is defined in this thesis as the "extent to which a product and its components are perceived as lasting over a long period of time without defects (product life)". Like for the other dimensions, this definition focuses on consumer perceptions instead of objective measurability that other authors have developed, e.g., by Hazen et al. (2017, p. 720) as the "amount of use before product failure or degradation severe enough to warrant obsolescence". Dur has been included previous PEF-related research (e.g., Agrawal et al., 2012; Jacobs & Hörisch, 2021; P. B. Jensen et al., 2021; Souri et al., 2018; J. J. Sun et al., 2021; Zamani et al., 2017).

Perceived repairability [Rep] is another aspect identified from the interviews and is closely related to Dur as it extends the useful life of a product after any possible defects. Similar to Dur, Hazen et al. (2017) have used reparability as dimension of PQ. Again, in the interviews, participants named Rep as important indicator of PEF. Note though that the relevance of Rep depends on the product category and is e.g., not applicable to food products. Based on participants' input and existing definitions in the literature (e.g., Hazen et al., 2017) it has been defined as the "extent to which a product and its components are perceived as effortless to bring back to a desirable state after defects". Rep has been investigated in previous research relating to PEF (e.g., Bracquené et al., 2021; Fachbach et al., 2022; Gobert et al., 2021; Makov & Fitzpatrick, 2021; Munten & Vanhamme, 2023; Nazlı, 2021; Rogers et al., 2021; Sonego et al., 2022; Van der Velden, 2021).

Perceived reuse- and recyclability [Rec] as a way to avoid trash was mentioned often. It gives consumers the feeling of being able to contribute to a product's PEF by bringing it back for recycling or reusing it themselves instead of disposing of the product or its packaging. The definition as the "extent to which a product and its components are perceived as having the potential to be used again after initial consumption for the same or other purposes rather than creating waste" reflects participants' consideration of disposal and waste. Like the other dimensions, also Rec played an important role in previous research in the area of PEF (e.g., Cooper & Gutowski, 2017; De Fano et al., 2022; Fogt Jacobsen et al., 2022; Greenwood et al., 2021; Y. Hu et al., 2022; Knoth et al., 2022; Otto et al., 2021; Roithner et al., 2022).

After having described and defined all dimensions in the previous paragraphs, Table 14 shows quotes from the interviews.

Table 14: Exemplary quotes from the interviews for each dimension of PEF.

Nat	• "It gets colored with natural substances and such, um natural colors, with plants, with stones, with all kinds of things. And then a product comes out there where I say, that is EF, that has grown, that is nature and it is also made with, with the people's hands." (17) • "That's EF, where the environment is not harmed by, by you don't apply chemicals which you yeah, which you actually don't need." (7) • "That perhaps not too many chemicals that are harmful to the environment are used" (15) • "Also means in manufacturing or growing, whatever I consume that there are no pesticides" (5)
Soc	• "I really strive for including social friendliness into being part of an EF, because if you say eco which is ecological and I mean human beings, as a part of ecosystem and ecological, ecological quality. Yeah, so for me, social friendless is hence also major part of that." (17) • "Is it produced… um ethical, like people or children work or whatever comes to your mind." (20) • "Also, like there's an element to me of social justice, so it's quite a broad understanding in my mind." (11) • "(…) how are the people treated over there and who is contracted, for example, what are the working conditions." (5)
Ani	• "The tests or anything, I know they don't test in animals anymore, which is good." (6) • "And still the animal was treated fairly." (5) • "So, also not from those mass things where the pigs have to live under such circumstances." (17) • "Or also if animals were harmed during production of the product." (16) • "And also, doesn't harm the living being let it be animals or, or creatures that live in the, in the sea." (19) • "That fulfill a certain level of, of 'Stallhaltung', of animal welfare." (12)
Cli	• "An EFP, for me, is something that is produced in a climate-neutral way." (9) • "Um just so that at the bottom line actually as few CO2 emissions as possible are caused - that would be EF for me." (8) • "My shallow understanding of it is, it is a product that doesn't contribute to the CO2 emission, ah the product doesn't contribute to the global warming" (19) • "I mean, first of all, CO2 emissions. But that's just, just one, just one factor. And that might be, there might be other… other figures or KPIs that I, that I, that's methane…" (4)

Res	• “Well, as I said, less resources. We have to see... water, so water is a limited product on the planet.” (20) • “For the production facilities of the product somehow not any forests are cut down, where you can build the production building.” (15) • “How much energy does it consume (…). That it's in a way, that the company found ways to minimize what is there really necessary. How much of the material is needed for creation?” (3) • “Also, the amount of water needed for production. (...) In meat production, or how much rainforest has to be cleared, for example.” (8)
Dur	• “And something that you can just wear for a really long time. Basically, something that's not a part of fast fashion, I think it's called, but something that can be worn, like for, for many years, for example.” (2) • “Yeah, so that it's lasting longer (…) with the attitude that it will, it will be of use for, for many years.” (3) • A product “which is long-lasting” (17) • “You always have to see how long do I use the product.” (20) • “So, the lifetime gets longer of the product, again it's more durable, it's more sustainable, because I don't have to throw it away that fast.” (10)
Rep	• “If a product ever breaks down, and you can just bring it somewhere and they repair it for you and then it's good as new (…).” (21) • A product “which can be repaired, really important” (17) • “How easily it can be repaired I think that's really important (…). Oh well, it's not that sustainable, is it if you like, the easier option is to replace it, instead of repair it.” (11) • “(…) also the life cycle duration of the product, is it repairable (…)” (4)
Rec	• “And also, what is really important for me personally, is what we will do with a product when you have used it. So, what about the disposal of the product. Can it be recycled or is it just burned somehow?” (10) • A product “which can be recycled or reused in a sensible way when it has to be disposed of.” (17) • “If they can be recycled or re… things can be reused, that would be the good aspect I would like.” (3) • “So, I would prefer if it were some recycled materials. Simply, to reuse the things also.” (8) • “And then, is a… like is easily recyclable, reusable etc.” (11)

4.3.2.2. Further considerations about PEF dimensions

Impact of specific dimensions. Taken together, the perceptions of the described dimensions form the overall PEF. However, the combination and impact of specific dimensions are influenced by additional factors: For example, different consumers can prioritize different dimensions (e.g., some prioritize Soc, while others are focusing on Rec). This seems to be especially connected to which dimensions they have more knowledge and seem more comfortable assessing (e.g., due to their education or work experience). Similarly, the type of product or which industry the product in question belongs to, can determine which dimensions have more influence. For example, Dur is perceived as far less important for food as for clothing or appliances while Nat is more important for food than for appliances.

Trade-offs between dimensions. Participants recognized that there are often trade-offs between several different dimensions. These trade-offs reflect the overall difficulty for consumers in assessing PEF. For example, participant 19 described they have a hard time estimating the PEF of glass (versus plastic) packaging

"because they don't contribute to the waste by using plastic and so on, so forth, but they still produce emissions through the transport and so on and so forth. Versus the plastic, yes, it is not EF because it's plastic, but it contributes less in the way it gets transported and so on and so forth, so it's very, very, very tricky".

Similarly, the trade-off between Cli and Res was mentioned by participant 4:

"So, if I buy a product, an EFP, then there might be impacts that I don't know that are not EF. So, um I considered it as EF because of reducing CO2 emissions, but on the other hand, it might and might need a lot more water to produce, for example, so...".

4.3.3. Understanding the correlations between PEF and PQ

In the following, the results in terms of PEF, PQ and impacting variables on their correlations are described. The elements are described one by one in the next sections and then the summary of all variables is shown afterwards.

One goal of this qualitative study was to understand consumers' self-reported reasons for correlations between PEF and PQ as previous research has shown that PEF can influence PQ in multiple ways (e.g., Borin et al., 2011; Frank & Brock, 2019; Hazen et al., 2011; Pancer et al., 2017). In line with previous research, participants voiced that PEF influences or can influence PQ either in a positive or in a negative way due to several reasons. When participants said they don't believe in any influence of PEF on PQ, there were no differing reasons but

agreement that PEF and PQ are separate things and PQ is determined by other factors. When participants reported an influence of PEF on PQ, there was a set of beliefs about influencing the PEF – PQ relationships. Those beliefs can be grouped into perceiver oriented, product oriented, and producer oriented, which are explained in the following sections.

4.3.3.1. Perceiver oriented beliefs

This section describes the constructs identified from the interviews as impacting PEF – PQ which relate to the feelings of the person assessing a product, i.e., the perceiver, thus being called perceiver oriented.

Anticipated warm glow [Wglow]. Several participants believe in higher PQ of PEFPs because they are looking for specific aspects to which they attribute positive emotions like being self-made, being hand-made, or being fair trade. Thus, they anticipate feeling good due to PEF which in turn influences their PQ. Those emotional contributions to PQ have also been found in other research (e.g., Davies & Gutsche, 2016; Habel et al., 2016). A participant believed re-used products can have higher PQ because of the story linked to these products. Another one believed in PQ of specific PEFPs because they have a connection to their own past, like soap bars. Others said that they get a better feeling from PEFPs and feel more connected. Concepts used in research that address the emotional aspects related to PEF are (anticipated) warm glow (Bezençon et al., 2020; Das et al., 2020) and anticipated conscience (Magnier et al., 2019). Hartmann et al. (2017, p. 45) defined warm glow as "an emotional construct of feelings of pleasure and satisfaction derived from the cognitive appraisal of contributing to the wellbeing of society, and, specifically, to environmental protection". Similarly, it has also been defined "as consumer's expectations regarding how the product will make him/her feel in an ethical sense" (Magnier et al., 2019, p. 85). Even though prosocial benefits have been included and measured in a very similar way by van Doorn and Verhoef (2011), they are more other-oriented than self-oriented. Specifically for Wglow, Bezençon et al. (2020) have shown its relevance relating to ethical attributes and product evaluations. The following paragraphs present exemplary quotes from the interviews about Wglow and the positive emotions.

"So, I am thinking about coffee or chocolate, so there I consume a lot for example from Gepard, who simply lead the fair trade though a bit and um… I always have the feeling that it is something much more special than some Lavazza coffee or some Milka chocolate or well… Or rather I simply have a better feeling like that and somehow it also tastes better for me." (Interview 16)

"If there is a human factor to it, or if there is, if I can connect to the human factor in that product then probably, I will perceive it as a higher quality item." (Interview 19)

"And it's also a consumer desire to be able to fulfill your consciousness and then feel happy about the fact I bought an organic apple today, instead of a conventional one." (Interview 12)

Beyond the findings on Wglow in the current qualitative work, this construct has a rather long history in research: "Warm glow of success" or "warm glow of giving" has been part of research at least since the 1970s (e.g., Isen, 1970), often associated with charitable giving and donations. While altruism and warm glow have been shown to both impact eco-friendly behavior, warm glow seems to be the main driver based on direct comparison in experiments (Hartmann et al., 2017). Warm glow and pure altruism differ in a way that warm glow is self-oriented while pure altruism is other-oriented (Bezençon et al., 2020), so that warm glow has been called "impure altruism" (Andreoni, 1990).

4.3.3.2. Product oriented beliefs

In this section, the constructs that relate to product attributes and were identified in the interviews as impacting PEF – PQ relationships are described.

Shared indicator beliefs [Indic]**.** One of the most commonly mentioned links between PEF and PQ were shared indicators, which means that the attributes or dimensions of PEF and PQ are perceived as being related. For example, a higher price or local origin can be signals for both PEF and PQ. On a more abstract level, being durable is not sufficient for PEF nor PQ on its own, but together with other relevant dimensions, it can increase both according to participants. Similarly, natural ingredients are perceived more EF and healthier, which is related to PQ (e.g., Binninger, 2017; De Pelsmacker & Janssens, 2007).

"Based on this sustainable creation, yeah so that it's lasting longer. So, products not of the food environment, but with the attitude that it will, yeah it will be of use for, for many years, so also goes, should go hand in hand with higher quality." (Interview 3)

"The more local the more quality so from an origin standpoint of view." (Interview 5)

"And if you have like no syn... synthetic ingredients in food and stuff of course it's healthier, it's better and because the, the, the, the holistic approach is different and yeah." (Interview 7)

Perceived product innovativeness [Innov]**.** A reason consumers believe in increasing PQ of PEFPs is that they believe in innovation and technological development of PEFPs. They

believe that *"there is also competition between those products, so they have to actually work, too"* (Interview 1). Thus, participants reasoned that the more PEFPs there are, the better quality they need to have to be interesting for consumers. Participants expected this trend for innovation and new EF technologies that increase PQ to continue growing. This belief is similar to beliefs about producers (e.g., having more experience), yet it is different in a way that it does not focus on a specific producer but rather looks on the product or product group as a whole.

"A lot is simply changing at the moment, also because of course the technology is quite different by now. And by now, thank goodness, sustainable products do really have a really great effect. So, they really can show great achievements." (Interview 13)

"A company (...) is using coffee grounds, which is being introduced into polyester fabrics, which is, which not only offers a sustainable aspect of the product itself but also offers added benefits. It reduces the... let's say smelliness of the product, and it also offers the additional benefit of UV protection, for example. And I guess, we will see more and more stuff like that happening." (Interview 18)

Innov has also been a central concept in previous research (e.g., Paparoidamis et al., 2019). However, the focus of Paparoidamis et al. (2019) was on innovation in terms of EF attributes rather than a broader view on product benefits as suggested by interview participants. A broader view on Innov was for example taken by Calantone et al. (2006), who found a positive relationship between Innov and product advantage (which they conceptualized very similar to PQ). Yet, the work of Calantone et al. (2006) was not done in the context of PEF.

Strength beliefs [Stren]**.** Even though participants said they believe in innovation and increasing technology of PEFPs, there remains a perceived gap regarding product characteristics: Especially strength related perceptions are lower for PEFPs, which can cause lower PQ for certain products, e.g., cleaning products. This connection between PEF and lower perceived strength has also been found in other research (Luchs et al., 2010; Mai et al., 2019), as has its impact on PQ (Skard et al., 2021).

"Um and I have… in fact, with Frosch cleaning products sometimes the feeling that they don't clean as well as that extreme Cilit Bang chemical mace." (Interview 16)

"And what I really do think that that us that the conventional ones are more aggressive with the ingredient or with the things they have inside." (Interview 5)

"For example, dishwashing detergent. There are just simply the conventional strong chemicals, they are just better in their cleaning power than the organic sustainable products, or my homemade dishwashing powder." (Interview 9)

4.3.3.3. Producer oriented beliefs

Constructs that relate to the producers of products and impact PEF – PQ relationships according to interview participants are described in this section.

Perceived producer eco-effort [Effort]. There is a common belief that producers should have a different approach to PEFPs than to CPs in a way that producers spend more effort on PEFPs. By spending more effort and putting more thought into a PEFP, consumers seem to believe that it relates to paying more attention to product quality. Participants associated PEFPs with higher PQ due to the belief that when producers look more closely at EF, those producers also look more closely at the whole production chain and introduce quality improvements. Part of this effect is known as *resource synergy beliefs* [RSB] in literature (Bodur et al., 2014; Das et al., 2020; Gupta & Sen, 2013). However, based on the interviews, this seems not to capture the full effect expressed by consumers, e.g., RSB (as operationalized by Gupta and Sen, 2013) are focusing more on socially responsible firm behavior in general rather than EF production as described in the interviews. Based on previous research, perceived eco-effort is defined in this thesis as consumers' evaluation of a firm's resources spent to create eco-friendly products (K. Lee et al., 2023; Morales, 2005). Exemplary interview quotes include:

"I have the feeling that the people that um… deal with the ingredients, um and where they are from, also attach more importance to quality." (Interview 16)

"Because I believe, if something is produced EF, or it is made EF, then the producer delves more into it and also cannot produce that quickly. And… and has things to, or has to retain from things and simply concern themselves more with the product." (Interview 14)

"And if you find for instance sustainable products, they, as I said, have by default a different approach and the approach is okay, they really, they really put thought into well how do we produce it, what do we put into it and not... They do not have, then the first thing is like, how can we cut costs. Like what is, what do we put in, what do we stand for, and then you can actually see and feel that as a consumer." (Interview 7)

However, when consumers believe that producers do simply switch their advertising approach for a product, then PQ can be affected negatively, for example because of suspicions that EF might be used to cover up other aspects. *"Because like that you make a product better than it actually is by saying it is EF, it is sustainable, whatever, and I wonder why do they think it is necessary to do this? And that's where quality plays a role for me, um if maybe it was done to, um because they do not have the best product and want to gain a good position anyways, to then say yes, but we are EF."* (Interview 14)

Perceived producer eco-expertise [Exper]. Consumers believed in increasing PQ of PEFPs because they thought producers have built-up more experience and expertise with PEFPs. Having more expertise means, for example, making less mistakes or having applied product improvements. They believe more focus of producers on EF and increased competition is helpful for increasing PQ. In previous literature, expertise describes "how competent and capable (…) the company in making and delivering the products they advertise" is (Newell & Goldsmith, 2001, p. 238). This concept of expertise has been adjusted as eco-expertise in this thesis, describing how competent and capable a company is in making and delivering eco-friendly products (instead of products in general).

As one participant put it: *"The quality yeah, based on... yeah based on much more experience throughout the year also increased there because when you start something new, or in a different way, then there is room for improvement and quite fast, they would go for it to really optimize the things".* (Interview 3)

In existing literature, the concept of producers' expertise is also related to credibility: Newell and Goldsmith (2001) have found that credibility perceptions of a given firm are formed from the perceived expertise and trust of consumers. Perceived credibility in turn has been related to PQ in previous research (e.g., P. Kumar & Polonsky, 2019).

Figure 11 visualizes the results as described in the previous sections: Product attributes (e.g., product origin or ingredients / materials) can impact the PEF dimensions (e.g., Nat or Soc), which impact the overall PEF of a product. The influence of PEF on PQ is partially mediated / moderated via the variables as described above. Overarching influences (e.g., product type / industry) can influence the relationships outlined at any stage within the model. Details about product attributes and overarching influences have been described in section 4.3.1, PEF dimensions in section 4.3.2 and the perceiver, product and producer oriented beliefs in section 4.3.3. Overall, this model describes a way how consumers' quality perceptions are impacted by product attributes via PEF dimensions, PEF, overarching influences, and specific perceiver oriented, product oriented and producer oriented beliefs.

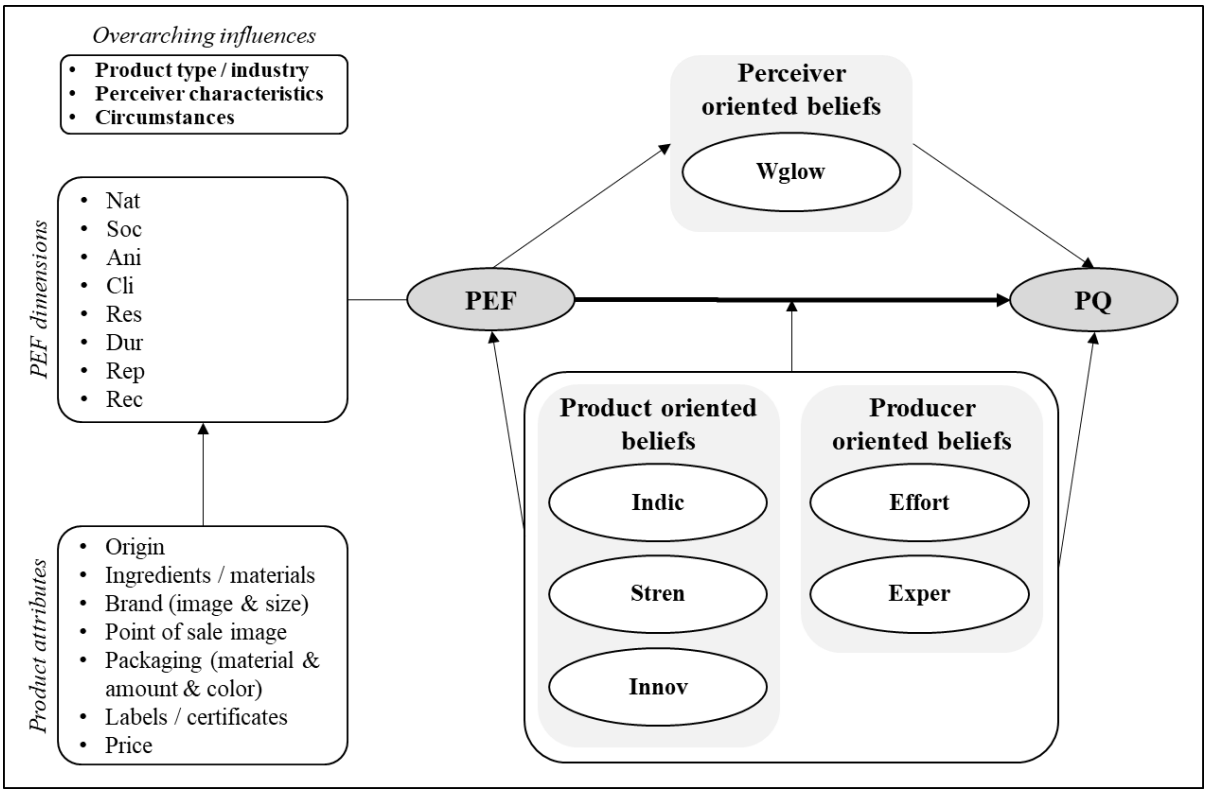


Figure 11: Visualization of the qualitative research results derived from the interviews.

This section has introduced Wglow, Indic, Stren, Innov, Effort and Exper as influencing variables on relationships between PEF and PQ according to consumers' self-reports. The effects of these variables can depend on product type, perceiver characteristics as well as circumstances as overarching influences. In the next section, the specific influence of each introduced belief is thoroughly reviewed and hypotheses deducted considering findings from previous research.

4.3.4. Hypotheses development

In this section, hypotheses for quantitative evaluation are formulated using the results outliend in the previous section and additional input from existing literature. Because PI is seen as another key variable within the PEF context (e.g., Das et al., 2020; Tong & Su, 2018; Zheng et al., 2018), the model for quantitative investigation will be extended to include PI (compared to the qualitative research results so far). The core of the model consists of relations between PEF, PQ and PI, which are described first. Then, the hypothesized effects of the perceiver oriented, product oriented and producer oriented beliefs identified from the qualitative research are presented. At the end of the section, the complete model under investigation is visualized and explained (Figure 12).

4.3.4.1. PEF, PQ, and PI

In this section, the hypothesized influences between PEF and PQ are described first and then extended to PI. These three constructs build the core of the overall model under investigation. PI is added as it is commonly used in marketing research to assess a consumer response (e.g., Dodds et al., 1991; Fuchs et al., 2015; Zeithaml, 1988) and also specifically in the context of PEF (e.g., Das et al., 2020; E. J. Lee et al., 2020; Lopes, Pinho, et al., 2024; Raj et al., 2023; Tymoshchuk et al., 2024) to such an extent that there are systematic reviews and meta analyses (e.g., K. Sharma et al., 2023; Singhal et al., 2019).

As shown in the systematic literature review within this thesis, the findings reported in literature on correlations between PEF and PQ vary strongly. In terms of direct effects of PEF on PQ, positive, negative, or no significant effects were found (see section 3.4.1). However, the author of this thesis expects a positive influence of PEF on PQ for several reasons: As indicated by interview participants and seen by comparing the articles from the systematic literature review, the correlations between PEF and PQ have changed over time, becoming more positive: Specifically, all articles included in the literature review that discussed any main effect between PEF and PQ and were published since 2022 found a positive correlation (see section 3.4.1). In addition, the correlations between PEF and PQ reported in literature vary between industries (see section 3.4.1). For food and clothing products, which are considered in the quantitative part of this thesis, the impact is often reported to be positive (e.g., Can, 2023; Chandra Pant et al., 2024; Farías, 2020; Hur, 2020; Magnier et al., 2019; Weisstein et al., 2024). Beyond that, previous research (e.g., Frank & Brock, 2019; Jäger & Weber, 2020; E. J. Lee et al., 2020; Mollenkopf et al., 2022) also has shown that referring to specific characteristics (rather than a broad claim) and describing multiple eco-friendly aspects (rather than only one), can influence PQ positively. The product descriptions applied in this thesis will contain several eco-friendly aspects specifically, thus supporting the expectation of a positive impact of PEF on PQ.

The impact of PQ on PI has been often reported in literature (Dodds et al., 1991) and a mediating effect between PEF and PI via PQ was found previously (e.g., Das et al., 2020). Based on the above described, the following hypotheses were formulated:

H1a. Perceived eco-friendliness (PEF) positively influences perceived quality (PQ).
H1b. Perceived quality (PQ) positively influences purchase intentions (PI).

4.3.4.2. Perceiver oriented

This section briefly introduces the hypotheses regarding the specific influences of perceiver oriented variables in the proposed model.

Wglow is defined as consumers' expectations regarding "feelings of pleasure and satisfaction derived from the cognitive appraisal of contributing to the wellbeing of society, and, specifically, to environmental protection" (Hartmann et al., 2017, p. 45). Wglow has been shown to influence various sustainable practices in previous research, e.g., using EF electricity (Hartmann et al., 2017; Van der Linden, 2018), choosing eco-labeled products (Bronnmann et al., 2021) or intending to reduce meat consumption (Taufik, 2018) and energy consumption (Taufik et al., 2016). Wglow also impacts product evaluations like PQ, perceived price fairness or WTP (Habel et al., 2016; P. A. L. D. Nunes & Schokkaert, 2003). The effect of Wglow is stronger when the behavior change is rather low-cost (Van der Linden, 2018), when consumers were exposed to visual advertisements that showed other images than pleasant nature (Hartmann & Apaolaza-Ibáñez, 2012) or when products have rather low previous consumer ratings (Bezençon et al., 2020). So far, most research relating to Wglow and EF has connected it to EF behaviors and sometimes products. Bezençon et al. (2020) have shown that Wglow can mediate the impact of PEF on PQ. Regarding PI, Iweala et al. (2019) found a positive impact of Wglow on PI for several different PEFPs, namely products labelled as fair trade, organic or carbon neutral. In line with these findings and previous research as described above (e.g., Bronnmann et al., 2021), Wglow is expected to increase PI. Thus, the following hypotheses were developed:

H2a. Perceived eco-friendliness (PEF) positively influences anticipated warm glow (Wglow).
H2b. Anticipated warm glow (Wglow) positively influences perceived quality (PQ).
H2c. Anticipated warm glow (Wglow) positively influences purchase intention (PI).

4.3.4.3. Product oriented

Within this section, the hypotheses for beliefs about product attributes (i.e., Indic, Stren, and Innov) are outlined as derived from the qualitative interviews and published research.

Indic describes the consumer belief that indicators or attributes for PEF and PQ overlap. For example, a relatively higher price can be linked to both PEF (Ryoo & Kim, 2023) and PQ (Bai et al., 2023; Yang et al., 2019). Similarly, for specific products, Dur can indicate PEF (J. J. Sun et al., 2021) as well as PQ (Hazen et al., 2017). Other examples given by interview

participants were local products or absence of synthetic ingredients. To date, no other research investigating Indic about PEF and PQ has been found in literature. Thus, for the formulation of the hypotheses, findings from the previous qualitative study are used. Since all products considered in the quantitative part of this thesis are described as EF and specifically mention shared indicators (e.g., natural ingredients / material), the researcher expects that Indic directly influences PQ. In addition, it is expected that Indic interacts with PEF in a way that the stronger Indic is, the stronger is the impact of PEF on PQ. Thus, the above and the insights from the interviews of the previous qualitative study led to the development of the following hypotheses:

H3a. Shared indicator beliefs (Indic) positively influence perceived quality (PQ).
H3b. Shared indicator beliefs (Indic) moderate the effect of perceived eco-friendliness (PEF) on perceived quality (PQ).

Stren describes the amount of strength (and similar attributes like effectiveness) that consumers associate with CPs over PEFPs. In previous research, Luchs et al. (2010) found what they call the *sustainability liability* of PEFPs: For specific product categories where strength is more desirable than gentleness to consumers, PEF can decrease preference for that product compared to a product with lower PEF. The authors explain that relation via lay theories, e.g., because ethicality and strength seem to be conflicting in sociocultural messaging. However, for products where gentleness is more valued, this negative impact of PEF is not expected to occur (Luchs et al., 2010). Similarly, Mai et al. (2019) draw upon such lay theories and further differentiate between explicit (reflective, conscious) and implicit (impulsive, subconscious) processes and show that both processes are relevant and interact with each other. Specifically, both explicit and implicit Stren have been shown to be able to impact sustainable product preferences (Mai et al., 2019). For the present study, it is expected that Stren not only directly influence PQ but also interact with PEF. For example, if a consumer expects a substantial difference in strength and perceives the product as high in PEF, the effect of PEF on PQ is expected to become less positive or even negative. Thus, explicit Stren are expected to impact PQ, and the following hypotheses were developed:

H4a. Strength beliefs (Stren) negatively influence perceived quality (PQ).
H4b. Strength beliefs (Stren) moderate the effect of perceived eco-friendliness (PEF) on perceived quality (PQ).

Innov has been shown to influence consumers' evaluations and behavior and specifically PQ. Innov is in this thesis defined as the degree of novelty of a product's features, functionality, and benefits, specifically relating to EF. This definition is in line with research that suggest that an innovative product should not only be novel but provide a meaningful update, e.g., bring increased value or utility (Y. Lee & Colarelli O'Connor, 2003; Shams et al., 2015). For the present research, this meaningful update is understood in terms of the product's EF. Thus, it is expected that if consumers perceive a product advertised via eco-friendly attributes as more innovate, they are more likely they also believe the product is more EF. If the product is perceived as more innovative and different from existing CPs, the PEF of that product is not or is less doubted (e.g., because greenwashing concerns are reduced).

Beyond that, according to the previous qualitative study, consumers expect advancement and thus higher quality from product innovations of PEFPs. This is in line with other studies that found positive correlations between Innov and PQ for product innovativeness (e.g., Boisvert & Khan, 2022), for service innovativeness (Senbabaoglu, 2017) and for brand innovativeness (Pappu & Quester, 2016). However, previous research also found a negative effect of innovativeness on PQ specifically in the context of EF (T. Park & Kim, 2024). One potential explanation for these seemingly opposing effects may be the nature of Innov in terms of incremental versus radical. Incremental innovations "refine and reinforce existing products and services (Subramaniam & Youndt, 2005, p. 452), while radical innovations "significantly transform existing products and services" (Subramaniam & Youndt, 2005, p. 452), so that existing designs or technologies may become obsolete (Chandy & Tellis, 2000). The researcher argues that in case of incremental Innov, consumers would infer higher PQ due to the refinement and improvement of existing products, while radical Innov may decrease PQ, e.g., due to lacking knowledge about the product (T. Park & Kim, 2024). The products / product descriptions used in the quantitative part of this thesis fall into the incremental Innov category. Based on the above, the following hypotheses were developed:

H5a. Perceived product innovativeness (Innov) positively influences perceived eco-friendliness (PEF).
H5b. Perceived product innovativeness (Innov) positively influences perceived quality (PQ).

4.3.4.4. Producer oriented

This section introduces the hypotheses related to Effort and Exper as influencing variables towards the relationships between PEF and PQ.

Effort. Previous research has shown that consumers overall reward companies that conduct extra effort (Morales, 2005). In the context of donations as CSR behaviors, higher Effort led to more positive attitudes toward the firm (Langan & Kumar, 2019). In the context of restaurants engaging in CSR activities, higher levels of Effort positively influence loyalty intentions (K. Lee et al., 2023), meaning for example how likely consumers are to say recommend the restaurant or consider it as first choice in the future. Additionally, for large brands (which are typically viewed more negatively in terms of PEF by consumers) increased Effort increases PI of their products via heightened authenticity (Wallach & Popovich, 2023). These findings complement interview results as presented in section 4.3.3.3 where participants connected higher Effort with higher PQ. However, the studies mentioned above have not investigated PEF as defined in this thesis and rather focused on donating time or money (Langan & Kumar, 2019), sponsoring broad CSR campaigns (K. Lee et al., 2023) or increasing effort as mitigation strategy for large firms (Wallach & Popovich, 2023). In addition to being connected with PQ, Effort is expected to be linked to PEF in the current study since products from companies who attribute more resources to EF should be perceived as more PEF. In accordance with the findings outlined above and extending previous research, the following was hypothesized:

H6a. Perceived producer eco-effort (Effort) positively influences perceived eco-friendliness (PEF).
H6b. Perceived producer eco-effort (Effort) positively influences perceived quality (PQ).

Exper. Specific investigations of producer expertise (and even more so of eco-expertise) seem to be scarce in literature even though expertise is essential to the overall producer brand image or reputation (Sen & Bhattacharya, 2001). According to Newell and Goldsmith (2001), perceived expertise and trustworthiness together define corporate credibility, which is more often subject in published research. Previous research has for example found positive relations between PQ of products and the perceived credibility of a specific brand or retailer (P. Kumar & Polonsky, 2019; P. F. Ng et al., 2014). Beyond looking at product quality, green brand credibility positively impacts green brand evaluation (P. Kumar et al., 2021). While these findings are based on the overall perceived credibility of a firm and not on Effort

specifically, they are in line with views from the interview participants within this thesis. In the interviews, participants reported impeding thoughts of greenwashing. It is reasoned that higher Exper counteracts these potential greenwashing perceptions of a specific producer since a company that is highly capable and experienced in producing PEFPs, would not need to engage in greenwashing behaviors. Overall, it is reasoned that the products of a producer perceived as having high expertise with eco-friendly products, are believed to be more EF and of higher PQ. Thus, the following hypotheses were formulated:

H7a. Perceived eco-expertise (Exper) positively influences perceived eco-friendliness (PEF).
H7b. Perceived eco-expertise (Exper) positively influences perceived quality (PQ).

After having introduced the individual hypotheses, Figure 12 summarizes the hypotheses described in this section. The model shows PEF, PQ and PI as focal constructs and the hypothesized influence of PEF on PQ and of PQ on PI, as well as the hypothesized influences of the other constructs. Wglow is expected to be influenced by PEF, while influencing PQ and PI directly. All other constructs included (Indic, Stren, Innov, Effort, Exper) are expected to influence PQ directly and either influence PEF directly or moderate the relationship between PEF and PQ.

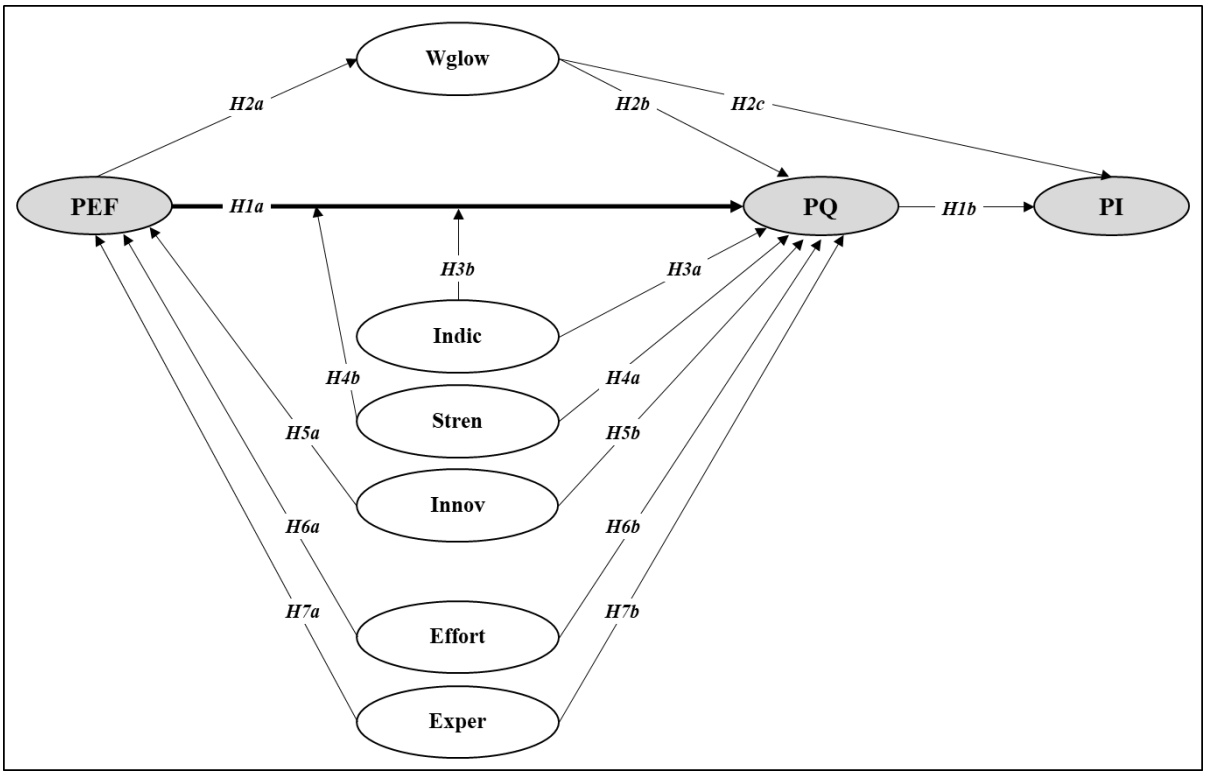


Figure 12: Theoretical model for quantitative investigation.

This section has further specified the expected relationships derived from the interviews in the context of existing literature. It has narrowed the focus of the overall qualitative research and built hypotheses for quantitative evaluation. In the next section, the overall qualitative results are discussed, before the quantitative research is presented in chapter 5.

4.4. Discussion of qualitative research

After the description of the results in the previous sections, the following sections show the implications arising for theory and practice and will discuss limitations of the current work, while connecting them to ideas for future research.

4.4.1. Theoretical implications

What is a PEFP, a product that is perceived as being eco-friendly? So far, marketing literature has not given an apparent answer that embraces the consumer perspective and covers all kinds of products at the same time. This work has used GT to conceptualize PEF and understand how consumer perceptions arise. To the best of the author's knowledge, this study is the first to define and conceptualize PEF holistically from this perspective. There have been others looking at specific parts, e.g., PEF packaging (A. T. Nguyen et al., 2020), or at specific industries, e.g., fashion (Henninger et al., 2016; Mukendi et al., 2020), but the overarching perspective was not taken. The recent and holistic view sets the basis for a mutual understanding of PEF between researchers and consumers. Research can build upon this understanding in multiple ways, e.g., by drawing on the knowledge of product attributes and PEF dimensions to operationalize or measure PEF, by further examining or extending the components of the model or by relating consumer perceptions to other measures of EF like life-cycle analysis.

Looking at operationalization and measurement, input from the interviews about PEF suggests PEF is seen as a multidimensional construct. This research has identified eight dimensions across a wide variety of consumer products, e.g., food, personal hygiene, clothing, and electronic products. The identified dimensions could be part of future quantitative research and be used to create a PEF scale that accurately reflects all its facets. When further investigating the dimensions of PEF, assessing the applicability of specific dimensions to specific product types / products could be an interesting next step for researchers to build on. Understanding the dimensions of PEF could for example lead to new measures of PEF or support researchers in operationalizing and choosing a suitable measure for PEF. Thus, the

research at hand extends the body of knowledge about operationalization and potential measurement of PEF, which can build the basis for future PEF research.

Looking at the components of the model, this work has made the first step in bringing them together and showing high-level correlations. Further examination of the specific correlations, e.g., between individual attributes or with other constructs can continue to improve the model. Regarding attributes, the research at hand is (according to the author's best knowledge) the first to assess them across product types / industries in such an extensive way. It cannot be finally determined if the list of identified attributes is exhaustive, for example because it might be changing over time. In general, the change over time was a recurring notion of participants' reflections on PEF, requiring research to be ongoing and critically re-assessing previous findings. Specifically, the relevance of coherent perceptions across attributes has been shown as well as the impact that an individual attribute can have. This can remind researchers to carefully consider all product attributes that consumers might perceive in their research. The current work has once more highlighted the importance of both intrinsic and extrinsic attributes, as well as the vital role of product packaging for the overall product perceptions.

The role of PEF for PQ has been increasingly considered in research (e.g., Blanco-Penedo et al., 2021; Das et al., 2020; Dixon & Mikolon, 2021), but so far with mixed results. The current qualitative work can bring more clarity into this field of research and opens the door for further investigation. For example, knowing that consumers find reasons for correlations between PEF and PQ in specific characteristics of the perceiver themselves, the product, and the producer can help to structure findings. The specific impact of the identified beliefs influencing PQ can be assessed more thoroughly in future research. In addition, findings from this research provide further support for some previous findings, e.g., that PEF is related to lower strength perceptions, which can then influence PQ for some products (Luchs et al., 2010; Mai et al., 2019).

On the other hand, new influencing factors that were (to the best of the author's knowledge) not mentioned in previous research about PEF – PQ relationships were identified, for example the producer oriented concept Exper or the product oriented concept Innov. Thus, the research at hand helps extend the body of knowledge regarding relationships between PEF and PQ.

At the same time, several moderators already recognized in literature for their impact on the PEF – PQ relationships were not mentioned by participants in this qualitative study. For example, the perceiver's ambiguity tolerance or the perceived intentionality of the EF aspects.

Reasons could be that consumers are not aware of these moderators or that their influence can be so complicated (e.g., depending in turn on many other variables) that consumers cannot think of examples or describe their impact. Thus, digging deeper into subconscious methods or potentially working with visual cues could further enhance existing knowledge in research.

Another topic for future research is to go one step deeper and analyze how different PEF dimensions could have different impact on PQ. For example, while consumers might associate a high perceived degree of Dur with higher PQ, they could associate a high degree of Res with lower PQ (e.g., due to perceived restrictions in production and decreased functionality).

4.4.2. Practical implications

For practitioners, this work provides valuable insights regarding the possibilities of understanding and shaping consumer perceptions. Policy makers can draw upon the findings regarding PEF to shape discussions and regulations, for example regarding labelling schemes or other product attributes. Using the presented insights about consumer perceptions, comparing them to objective EF measures and deducting action plans on how to align the two can help promoting objectively eco-friendly products. Additionally, the impact of consumer education and knowledge has again been shown in this research, thus highlighting the need for suitable campaigns. First topics for knowledge creation have been identified (e.g., product attributes and PEF dimensions). How specifically they can be addressed most effectively and efficiently needs further investigation.

Not only policy makers but also managers of varying functions, e.g., marketing or sustainability, can leverage the findings from this research. They can easily gain an overview of which attributes and dimensions are important for consumers regarding PEF. Based on that, managers can develop strategies to shape perceptions about their eco-friendly products and try to counteract greenwashing perceptions, e.g., how to position their brand, where to sell their products, and which cues to show on the product. Specifically for brand management, the importance of a holistic EF image rather than singular EF product lines and the importance of considering the retailer image have been highlighted. Further, knowing where consumers are most susceptible to greenwashing can help to raise awareness against it and properly address greenwashing issues to create a strengthened brand image and consumer trust. Even though some of the results of this qualitative research could be misused to engage in misleading marketing and greenwashing, the author does not want to encourage these behaviors but rather provide transparency and trusts in the responsible use of the findings.

Similarly, by knowing the reasons why consumers have higher or lower PQ for PEFPs, managers can address those topics in their marketing activities. For example, campaigns can tell consumers specifically about the thoughts beyond EF that went into the product, how it is innovative or about the producers' motives and experiences with specific products. They can also appeal to the positive emotions and Wglow that customers can obtain by purchasing or using PEFPs to increase these products' PQ. These are some reasons why consumers believe in an influence of PEF on PQ. Doing so will likely convince more consumers to try those products and then potentially be convinced by the way they work after having their own experiences. Thus, it can help to decrease prejudices against PEFPs.

4.4.3. Limitations of conducted qualitative research

As with all research, this study has some limitations. First, even though a sample size of 22 is in a typical range according to Zeithaml et al. (2020) and Thomson (2010), and theoretical saturation was given, there might be other opinions and beliefs that could not be detected in this qualitative interview-based research. To address this and continue to refine the theoretical model across diverse contexts and with a broad population, the qualitative study will be followed-up with quantitative research.

Second, this research has a geographical and age-related focus that needs to be considered and potentially re-assessed. The focus of this research is put on European countries, specifically Germany. In the 2020 Environmental Performance Index, European countries were on top of the list (Wendling et al., 2020) and European consumers see climate change as biggest threat among several (Poushter & Huang, 2020). In addition, participants in the interviews believed Germans to be rather well informed about EF topics. Thus, results might look different for other countries or regions, e.g., developing countries. Similarly, the study considered only adult consumers meaning that results specifically for children and adolescents could look different and be investigated in future research. Thus, future research could expand this research (e.g., geographically) and review the developed theory after some time.

Third, all results from this research are at least to some degree subject to being consciously experienced by consumers as they needed to be aware of their thoughts, beliefs, feelings, and perceptions to be able to discuss them. To build new theory, this method was purposefully chosen, and participants were thoughtfully challenged in their answer via follow-up and probing questions to uncover the actual underlying phenomena. Nevertheless, the role of unconscious beliefs could be explored in future research to complement existing findings.

Fourth, this research has focused on the impact of PEF on PQ, without addressing the potential for influence in the other direction: From PQ to PEF. Thus, another topic for future research is the investigation of the impact of PQ on PEF. For example, a few participants reported perceiving products with higher PQ as more EF and specifically drew the direction of influence in that way. Especially shared indicators like Dur can have the potential to work in both directions (J. J. Sun et al., 2021), but there could be other additional effects.

Fifth, the findings from this qualitative work represent consumer views at a specific moment in time which may change in the future. Findings from GT are commonly temporally constrained, meaning they might change over time and the topic needs to be re-evaluated eventually (Goulding, 2002). While this may be a limitation of consumer research in general, it is especially applicable in the context of EF as indicated by interview participants who discussed changing views in the past few years. This is also in line with the relatively recent, developing and changing attention to this topic in research and practice. Thus, future research should re-inspect and adjust or extend the findings further.

4.5. Summary of qualitative research

This qualitative study has developed a model for consumers' PEF, PQ, and covariates by applying GT. Findings from the analyses of 22 intensive interviews imply that the consumer definition of PEF is similar to existing academic definitions with only small refinements applied. This study has also identified an extensive list of product attributes that shape consumer perceptions of PEF (e.g., origin, ingredients, brand, point of sale image, packaging labels and price), while additional influences of perceiver characteristics, circumstances, and product type / industry also play a role. In addition, interview findings imply that consumers often understand PEF as multidimensional. Eight dimensions (Nat, Soc, Ani, Cli, Res, Dur, Rep, and Rec) that together create the overall EF perception were identified, defined, and explained with examples in this article. Lastly, this study has also found that consumers attribute the impact of PEF on PQ on 1) themselves as perceivers, 2) perceived differences of the product itself and 3) perceived differences regarding the producers. Within the first category, Wglow was identified as mediator between PEF and PQ. Within the second category, Indic and Stren were identified as impacting the PEF – PQ relationships as moderators while impacting PQ directly. Innov was identified as directly impacting both PEF and PQ. Within the third category, Effort and Exper were identified as impacting both PEF and PQ directly.

This study extends existing research in several ways, for example by sharpening the definition of PEFPs, by proposing a novel multidimensional conceptualization of PEF and by identifying novel constructs impacting the relationships between PEF and PQ (e.g., Exper). This qualitative work also shows avenues for future research, including further investigation of the proposed PEF dimensions, the quantitative validation of the developed model, the extension to consumers with differing cultural backgrounds (especially regarding awareness and knowledge about EF), or the complementation of this research with subconscious measures to broaden the methodological spectrum applied.

In the next chapter, the quantitative research which directly builds onto the qualitative research is described. The quantitative research will test and start validating the theoretical model as identified from the qualitative research in this thesis and explained in the past sections.

5. Quantitative research

This chapter describes the quantitative study conducted within this thesis, starting by the specific aim and research questions. It then explains the methods of quantitative research (including pretests, data measurement, data collection, sample and analyses conducted) to test the hypotheses derived from the previous qualitative research, and then reports the respective results. Lastly, it discusses and summarizes the quantitative study.

5.1. Aim and research questions

This quantitative study aims to further examine the PEF dimensions and the constructs identified in the qualitative study regarding their specific impact relating to PEF, PQ and PI. Using a survey design and statistical analyses, this approach enables the testing of theoretical relationships under varying conditions, such as the unidimensional versus multidimensional conceptualization of PEF and across different product categories, and to generalize findings for a broader population. In other words, quantitative assessment is applied to further refine the model as developed via GT in the previous chapter. In total, a more thorough and nuanced understanding will overall be created by combining both qualitative and quantitative approaches in this thesis. This aligns with established marketing research practices, where quantitative methods are frequently used to analyze construct validity and assess theoretical models. Furthermore, this approach contributes to more nuanced insights by accounting for potential moderating variables beyond the hypothesized effects, e.g., demographic characteristics. Overall, this allows to contribute theoretical and managerial insights about product evaluations and (eco-friendly) consumer behavior.

In line with the aim outlined above, the following research questions guide the quantitative work in this thesis:

Which of the proposed constructs influence the relations between perceived eco-friendliness (PEF) and perceived quality (PQ)?

b) How do these influences vary across selected product categories?

a) How do these influences differ between the unidimensional and multidimensional approaches to perceived eco-friendliness (PEF)?

The remainder of this chapter first explains the quantitative research methods applied in section 5.2. This includes detailed information about the questionnaire design, measures and control variables. Section 5.3 describes the data collection process and provides an introduction

to the analyses applied. Afterwards, the results of the structural equation models are presented in section 5.4, followed by a discussion of the results (section 5.5) and finally a brief summary of the quantitative research (section 5.6).

5.2. Quantitative research methods

In the previous chapters and sections, the model evolving around PEF and PQ has been developed and specific hypotheses to be assessed have been formulated. The following sections explain the quantitative methods and procedures applied in the current research. First, two pretests are described, then the data measurement (including a general methodological background, the questionnaire design, and the specific measures and control variables assessed) and finally the data analysis which introduces the specific methods (i.e., descriptive analyses, factor analysis, regression and structural equation modelling) is explained.

5.2.1. Pretests

This section explains the pretests that were conducted to prepare the main study. Pretest means in this thesis a small-scale study conducted before the main study with the purpose of refining the content of the main study. In the current work, two pretests were conducted before the main study to identify suitable products for the study (pretest 1) and to ensure that the questionnaire is clear and understandable with proper measures (pretest 2). The procedures and key outcomes of the pretests are explained in the following sections.

5.2.1.1. Pretest 1

The first pretest contained a total of 16 different products with the goal of identifying the best suitable products to be used in the main study. The pretest was conducted as online survey using the same platform as the main study, namely Unipark (2024). One of four different conditions was selected by the survey software for each participant – mostly random but still ensuring that conditions reached a similar number of participants overall. Each condition contained four products to assess based on groups of related products: Food, Cosmetic / Hygiene, Clothing and Durables / Machines. The initial list of 16 products was developed based on theoretical and practical considerations: First, products should be common and known to most consumers. Second, the 16 initial products should be diverse overall and together ensure that each of the eight PEF dimensions identified via the previous interviews can apply. Third, most of the products chosen were previously used in similar research (e.g., Magnier et al., 2019; Newman et al., 2014; J. C. Nunes et al., 2021; Paparoidamis & Tran, 2019; Ryoo & Kim, 2023;

Skard et al., 2021; Usrey et al., 2020; Van Doorn et al., 2020; Weisstein et al., 2024). Fourth, product examples provided by participants in the previous interviews were considered as input (which led for example to include meat substitutes as they were described as high PEF but varying PQ).

The questionnaire was presented on a single page. First, participants read the definition of PEF and were asked to envision four different products, e.g., eco-friendly dish soap, eco-friendly laundry detergent, eco-friendly body lotion and eco-friendly toilet cleaner (in the condition with Cosmetic / Hygiene products). Then, they rated PQ and PEF of each product (using a 5-point Likert-type scale about the level of agreement to specific statements), followed by single-item measures for all eight PEF dimensions (4-point Likert-type scale about the level of importance of each dimension). A differing number of answer options was chosen to make it apparent that the meaning of the answer options is different. For this pre-test, single item measures for the PEF dimensions were chosen to limit the complexity for the participants and decrease mistakes (e.g., due to fatigue or inattentiveness) or drop-outs. Because the purpose of the pretest was to select products for the main study and not to test hypotheses, single-item measures seemed a better fit considering the benefits and drawbacks even though multi-item measures are often preferred (Diamantopoulos et al., 2012). Additionally, some authors of reputable publications use single-item measures in their studies (e.g., Bodur et al., 2016; Weisstein et al., 2024).

The questionnaire was distributed via the personal network (including snowballing) of the researcher in August 2024. In total, 44 participants completed the questionnaire, but one participant was excluded due to a response time of less than 3 minutes (average response time of 12 minutes across all participants), leading to a final sample size of 43.

The selection of the final products for the main study was based on a pre-specified set of criteria, mostly regarding the PEF dimensions. These criteria included the number of relevant PEF dimensions per product (at least three dimensions rated as important), representation of all PEF dimensions across the selected products and not selecting two products from the same product group (i.e., from a different scenario); and other factors like participants' comments in the questionnaire. Based on the criteria described, the following products were selected for the main study: Chocolate bar and Sweater.

5.2.1.2. Pretest 2

The objective of pretest 2 was to ensure all questions are understood and the flow of the survey is clear. This pretest happened in two phases: First, ten participants filled a preliminary version of the survey in the presence of the researcher in so called *think aloud sessions,* as commonly used in literature (e.g., Hunt et al., 1982; Kaczorowska et al., 2021; Osburg et al., 2019). In those, participants were asked to voice their doubts, uncertainties and thoughts when filling out the survey to identify any potential misunderstanding or challenges. Four of the participants followed this procedure for both the German and the English version (see section 5.2.2.2 for more information about the design and language choice) and were asked if they detected any relevant differences in understanding of the questions and statements.

Second, 40 answers were collected via self-administered testing of the online survey. Participants received the preliminary online questionnaire and were asked to answer truthfully and provide feedback in case they notice anything unexpected, unclear, or disturbing. Participants could choose if they wanted to fill in the survey in German or English. This step was taken to get a first glimpse into the psychometric properties, get an indication of duration to fill the questionnaire and uncover any potentially remaining challenges (like unclear wording, typos, etc.).

Based on participant feedback in both stages of pretest 2, minor adjustments to the questionnaire were implemented (especially regarding wording of instructions, wording of demographic questions and answer options).

5.2.2. Data measurement

This section first introduces key terminology, content, and implications from measurement theory. It then explains the measurement of the main study in detail by laying out the general questionnaire design, introducing the specific measures and items included and finally presenting which and how control variables were included.

5.2.2.1. General methodological background

This section is based on measurement theory. It begins by clarifying key terminology such as latent variable, indicator, or reflective / formative measure. Afterwards, quality of measurement is addressed, introducing reliability and validity and how they are considered in this thesis.

In social sciences, including marketing research, (theoretical) constructs of interest are often not directly observable but need to be operationalized, i.e., made measurable (O. Jensen,

2004). Constructs or unobservable attributes that are described via observed variables are commonly called *latent variables* (Borsboom et al., 2003), while the observed variables are commonly called *indicators*. In a questionnaire, an indicator is also called an *item* (i.e., a single statement or question) and a specific set of items or indicators that together operationalize a latent variable or theoretical construct is called a *measure* (American Psychological Association, 2018b).

The relationships between indicators and their underlying latent variable can be defined in two different ways: As *reflective or formative constructs* (Diamantopoulos & Winklhofer, 2001). In a reflective definition, "indicators are seen as functions of the latent variable, whereby changes in the latent variable are reflected (i.e., manifested) in changes in the observable indicators" (Diamantopoulos & Siguaw, 2006, p. 263). According to the same authors, the indicators of a construct defined in a formative way cause a latent variable and thus are not interchangeable. Additional differences between the two and implications of a formative perspective are discussed by Diamantopoulos and Winklhofer (2001). For the research at hand, all latent variables are operationalized and defined in a reflective way both due to theoretical considerations and in line with previous research from which measurement scales have been adopted. The applied measures will be presented later in this chapter (section 5.2.2.3).

Specifically for PEF, the operationalization as reflectively defined latent variable implies that different indicators can be chosen for different products because several indicators can be interchangeable from the reflective perspective. According to the previous qualitative research, PEF can be operationalized via up to eight different dimensions which are assumed to differ between products / industries. Due to the reflective definition, these sub-sets of indicators can be interchanged while still defining the same latent construct, in this case PEF.

After having defined the latent variables as reflective, another important aspect is understanding how to assess the quality of measurement. To do that, *reliability* and *validity* are key concepts a researcher considers (next to objectivity) (Bühner, 2011). Highly summarized, all multi-item measurement scales typically contain some error. This error can be systematic or unsystematic (in other words random). If the measure has very little random error, it is highly reliable; and if it has very little random and systematic error, it is highly valid (Drost, 2011). In more lay terms, high validity means that a measurement scale measures the construct it is supposed to measure; and high reliability means that a measurement scale measures well and consistently (regardless of what it measures). Thus, high reliability is necessary but not

sufficient to reach high validity. The nuances of both these concepts for assessing measurement quality will be explained in the following paragraphs.

Reliability "can be defined conceptually as the correlation between a measure and itself" (Peter, 1981, p. 136). A highly reliable measure means that random errors are minimized so that measurement is consistent (e.g., over time). There are different ways how to empirically assess reliability (Bühner, 2011):

- *Internal consistency* uses responses from one single measurement and sample and splits them into sub-tests. Two common variants exist: First, split-half reliability where a test is split in half and then the correlation between the halves is used to estimate reliability. Second, and most commonly used, internal consistency is estimated via (co-)variances of all individual items. This second type will be referred to in the remainder of this work when using the term internal consistency. Analyses regarding internal consistency should only be applied for a specific construct if it is reflective and unidimensional.
- *Parallel-test reliability* uses two parallel forms of measurement for a construct within the same sample (at the same time). The reliability is then estimated as the correlation between the scores of the two parallel measurements.
- *Test-retest reliability* uses the same measurement with the same sample after a certain amount of time has passed. The reliability is then estimated as correlation between the scores in the first and the second measurement.

Malik et al. (2021) suggest using Cronbach alpha (Cronbach, 1951) to assess internal consistency reliability, and factor loadings to assess indicator reliability of a latent construct. Following this advice and due to the complexity and cost of assessing test-retest or parallel-test reliability, reliability in the current study will be assessed based on Cronbach alpha (and factor loadings). Only for PEF, parallel-test reliability will be investigated by measuring PEF with a unidimensional 3-item measure (see section 5.2.2.3) and in addition measuring PEF via multiple dimensions. Factor loadings will also be assessed in the context of a factor analysis (see section 5.2.3.1 for more details). For internal consistency of the multi-item scales, Cronbach alpha will be complemented by McDonald omega (R. P. McDonald, 1999) in light of criticism towards Cronbach alpha (Hayes & Coutts, 2020). Despite the criticism, Cronbach alpha is one of the most commonly used criteria to assess internal consistency (Dunn et al., 2014; Hayes & Coutts, 2020). In order for Cronbach alpha to be a good indicator of reliability, several assumptions need to be met – which is often not the case (Dunn et al., 2014). However,

as long as the item error terms are uncorrelated, alpha can be understood as "a lower bound estimate of reliability" (E. Cho & Kim, 2015, p. 213). McDonald omega further relaxes assumptions needed for alpha to be a good estimator of reliability and is thus typically more precise (Hayes & Coutts, 2020). In addition to using McDonald omega, a confidence interval of the McDonald omega point value (obtained via bootstrapping) is assessed to achieve "a more accurate degree of confidence in the consistency of the administration of a scale" (Dunn et al., 2014, p. 406). Specifically for the *structural equation model* [SEM] analyses, rhoC will be reported instead of McDonald omega because it does not assume equal indicator loadings (Hair et al., 2021). RhoC is very commonly used in *partial least square structural equation model* [PLS SEM] analyses (this method is explained in section 5.2.3.3). In terms of recommended thresholds of reliability coefficients, there is no hard rule but often values > 0.7 are considered suitable, with some authors cautioning about values > 0.95 as that would indicate high redundancy and inefficiency (Taber, 2018; Wong et al., 2020).

Validity can be defined "as the correlation between a measure and the true underlying variable" (Heise & Bohrnstedt, 1970, p. 123). Thus, it indicates how well a measure can capture the construct it is supposed to assess and can be further broken down into several sub-types that together ensure the overall conceptual fit of a measure (Peter, 1981):

- *Convergent validity* is given if different measures of the same construct create very similar results (Eid et al., 2011). Convergent validity is rather closely related to internal consistency and can be assessed via the *average variance extracted* [AVE] estimate (Malik et al., 2021), which will be used in this thesis, particularly in relation to the factor analyses and SEM. Values of ≥ 0.5 are typically recommended (Cheung et al., 2024). In addition, factor loadings can indicate convergent validity as they describe the relationship between an individual item and the factor. A threshold of ≥ 0.4 is commonly applied in research (Deo & Prasad, 2024).
- *Discriminant validity* is given if measurements of different constructs – that are theoretically not supposed to be related – actually are not or minimally related (Eid et al., 2011). Historically, the *Fornell-Larcker criterion* [FLC] (Fornell & Larcker, 1981) was commonly used to assess discriminant validity; more recently though the *heterotrait-monotrait ratio of correlations* [HTMT] (Henseler et al., 2015) has become the recommended method (Voorhees et al., 2016). In this work, both FLC and HTMT will be assessed as this "should be the standard for publication in marketing" (Voorhees et al., 2016, p. 119). For FLC, it is recommended that all

individual squared correlations between one construct with the other constructs of the study should be lower than the AVE of a given construct (Voorhees et al., 2016). For HTMT, the authors suggest using a cutoff value of ≤ 0.85 or 0.90, with higher values indicating a lack of discriminant validity (Henseler et al., 2015).

- *Nomological validity* "is primarily 'external' and entails investigating both the theoretical relationship between different constructs and the empirical relationship between measures of those different constructs" (Peter, 1981, p. 135). It locates a construct within a theoretical model and describes its relationships with other constructs. Borsboom (2004) argue that nomological validity cannot be sensibly assessed for psychological constructs due to lacking detailed theory. Similarly, nomological validity will not be specifically assessed in this thesis. A confirmation of hypotheses may provide first evidence for nomological validity, however it cannot be comprehensively assessed within a single study.

Beyond validity of a construct, validity of a whole study is also relevant to understand how the findings may be further used. In terms of validity of a study, internal validity means that findings can be unambiguously attributed to actual relationships between variables without systematic interferences, and external validity means that findings from that study can be generalized and transferred to e.g., other samples, places or times (Eid et al., 2011). For the present work, internal validity is supported by including several potential control variables to ensure that their potential systematic interferences can be accounted for (see section 5.2.2.4 for all control variables). In terms of external study validity, the study is not carried out in a laboratory but is accessible to almost anyone and from anywhere with internet access. In addition, the product descriptions used were aimed to be as realistic as possible. Table 15 summarizes the criteria and thresholds applied in this research to estimate reliability, convergent validity, and discriminant validity.

Table 15: Key criteria and thresholds applied to assess reliability and validity. Source: Adapted from Cheung et al. (2024) and Ringle et al. (2023).

Criteria	Threshold	Applied to…
Reliability		
Cronbach alpha	≥ 0.7 (and < 0.95)	Factor analyses, Structural equation modelling
Mc Donald's omega	≥ 0.7 (and < 0.95)	Factor analyses
rhoC	≥ 0.7 (and < 0.95)	Structural equation modelling
Convergent validity		
Factor loadings	≥ 0.4	Factor analyses, Structural equation modelling
AVE	≥ 0.5	Factor analyses, Structural equation modelling
Discriminant validity		
FLC	$AVE_x > r^2_{xy}$	Structural equation modelling
HTMT	≤ 0.85 (or < 0.90)	Structural equation modelling

Beyond assessing reliability and validity as described in the paragraphs above, additional attention was paid in this quantitative study to support its *objectivity*, *reliability*, and *internal and external validity* (not specifically related to a certain measure) - for example via randomizing the order of statements / answer options or asking participants to act in their natural role as consumers assessing common "everyday" products. After having introduced key concepts and quality criteria for latent variable measurement, the next section will describe the questionnaire design and measures applied in this thesis in more detail.

5.2.2.2. Questionnaire design

This section describes the structure, design, and content of the questionnaire, including its general design, language, visual design and order of content, as well as the scenarios and product descriptions used.

General design. Data measurement occurred via a structured online questionnaire. Participants first learned about the anonymous and voluntary nature of the study, gave their informed consent, chose the language for the survey, and then read a product description before answering the questions. Only closed questions with pre-determined answer options were used in line with the aim of quantitatively testing a developed model (as opposed to generating new ideas via open questions). One exception was made at the end of the survey where participants

had the option to provide feedback or ask questions via an open text field. All answer scales (apart from the demographic questions) are in the format of Likert-type scales as is common practice in this type of research (e.g., Atkinson & Rosenthal, 2014; Das et al., 2020; Frank & Brock, 2019). The survey did not include answer options like "Do not know" or "Prefer not to answer" because research has shown that those options can lead to more missing data compared to not offering such answers (Kmetty & Stefkovics, 2021). However, at the same time, forcing participants to answer each question leads to higher drop-out rates (Kmetty & Stefkovics, 2021) and may raise ethical challenges (Tourangeau et al., 2013). Thus, participants were informed in the beginning that they can drop out at any time without facing any consequences.

Questionnaire language. To reach a broad audience, the questionnaire was available in English and in German. After reading the initial page with instructions in both languages, participants chose in which language to progress. All original measurement scales were available in English. They were translated into German by the researcher and then translated back to English using Google Translate (Google, n.d.). The results were in addition reviewed by two bilingual speakers (native German, C-level English) from the personal network of the researcher (who communicate in both English and German in business and in everyday life). Where necessary, adjustments to the German version were made. In addition, wording feedback from participants of pretest 2 was considered in both German and English.

Visual design and order. Regarding the visual design of the survey, the questions were distributed across several pages rather than shown on one long page requiring to scroll down. As Tourangeau et al. (2013) describe, this is more common and has several advantages, e.g., allowing the researcher to control the order of questions more and permitting error checks. Additionally, paging design can take less time to complete and even decrease item nonresponse (Peytchev et al., 2006). Another aspect of survey design involved the order of questions and response options within questions: Order effects have been long documented in literature (Strack, 1992). Measures for the dependent variables (PQ and PI) are placed at the beginning as these are rather straightforward to answer (and are thus motivating to continue the questionnaire) and to avoid participants guessing the purpose of the study over the course of the questionnaire and then adjusting their responses to the dependent variables. For questions with several items, the order in which individual items were presented was randomized between participants by the survey software to avoid any confounding effects of item order. This approach is in line with recommendations by Krosnick and Presser (2009). Within Likert-type scales, the order from left to right remains the same, i.e., from "fully disagree" to "fully agree".

Scenarios and product descriptions. At the beginning of the questionnaire (after providing consent and choosing the language), participants were presented with a specific product scenario and asked to read it in detail. The survey software randomized which product scenario (out of two options) was shown to a given participant but also ensured that the number of participants per scenario did not differ strongly. Similar to other studies (e.g., Borin et al., 2011; Steinhart et al., 2013), scenarios were used where participants were asked to envision a shopping situation or read a product description before answering questions. The product descriptions were intentionally kept short and concise to decrease the likelihood of participants rushing over the description and missing details. Like in other studies, only verbal descriptions were used without images to avoid potential aesthetic design influences (e.g., Munten & Vanhamme, 2023). Product descriptions were intended to be highly realistic. Thus, descriptions of the eco-friendly aspects of the products were informed by the comments and feedback from pretest 1 as well as real product descriptions found via Google search. For each product, typically advertised characteristics (e.g., ingredients / origin) were described in a way to create eco-friendliness perceptions, especially for the respective dimensions of PEF considered as important in pretest 1 for each product. Additionally, both core and peripheral product characteristics were described as eco-friendly because they could have different impact (Skard et al., 2021). As the product description is framed as product advertisement, a broad claim like "better for the environment" was not included to avoid raising greenwashing concerns, even though such broad claims have been successfully used in previous research (e.g., Newman et al., 2014). Specific information about the producer (e.g., in terms of expertise and their effort related to EF) was not included as this is not commonly part of product advertisements and should also be inferred from the description of eco-friendly attributes. Similarly, specific information about product innovativeness was not included beyond the description of the eco-friendly attributes in line with the operationalization of innovativeness in terms of incremental innovativeness related to EF of the product. Beyond the description of the eco-friendly aspects, each description contained information unrelated to EF (e.g., for the Chocolate bar: "Choose the flavor you like most, for example whole almond, salty caramel or dark (80%) chocolate" to make the scenario more realistic and mask the study objective. At the end of the product description, participants were reminded that there are no right or wrong answers (Borin et al., 2011). See Appendix A for the detailed product descriptions.

5.2.2.3. Measures

This section describes the specific measures applied in the quantitative research and how they were selected or constructed. *Measure* means in this thesis "an item or set of items that provides an indication of the quantity or nature of the phenomenon under study" (American Psychological Association, 2018b), with an item being an individual statement that participants answer by expressing their level of agreement. For all constructs, multi-item measures (i.e., measures containing more than one item) were chosen as they are typically preferred over single-item measures, e.g., because common assessment of measure reliability can only be applied to multi-item measures and predictive validity tends to be better for multi-item measures (Diamantopoulos et al., 2012). Established measures from other authors have been taken wherever those could be identified in literature. Search for measures in literature was conducted in the following way: First, articles included in the systematic literature review and in the qualitative study were searched for applicable measures. Afterwards, keyword search on EBSCOhost and Google Scholar was conducted, using the construct name and closely related wording (e.g., "animal friendliness" or "animal friendly" or "animal welfare").

When multiple measures were available in the literature for one construct, selection for this survey was conducted based on five criteria (sorted by decreasing importance): 1) Being a multi-item measure (at least three items); 2) Being coherent with the content and wording of interview participants; 3) Being suitable to all products under investigation; 4) Having been applied in a similar context involving PEF and / or PQ; 5) Being used in a journal ranked B or higher in VHB-JOURQUAL3. Specifically for PQ, numerous different measures are available in the context of PEFPs. However, many of those measures are specific to a certain product group and do not fit all products under investigation: For example, they include items about efficiency, reliability or workmanship (which do not apply for the Chocolate bar or food, clothing, etc. in general), or healthiness or tastiness (which do not apply to the Sweater or any kind of clothing, appliances, etc.). Thus, for PQ, the measure of Stolz & Bautista (2018) was initially chosen as universally applicable, three-item measure that was closely related to how interview participants had described PQ (including being adapted to their needs). However, in the pretests, it was mentioned that the term "high standard product" might be difficult to understand so that an additional item was added from Konuk (2018), whose measure already shared an item with Stolz & Bautista (2015). Table 16 shows the applicability of each criterion to each measure. Note that if two articles are cited as source, the measure applied is a

combination of the items applied in each. This was done a few times to ensure that each measure consists of at least three items.

For constructs where no applicable measures could be identified in the literature (i.e., Indic and Rec), own items were developed based on the content and wording participants used in the interviews. These self-developed items were discussed first with four previous interview participants and then with participants of pretest 2 to ensure they capture the essence of the construct under investigation and are understandable. Slight adjustments in terms of wording were incorporated in the version included in the final questionnaire.

Table 16: Application of criteria for each selected measure.
Notes: x → fully applicable; o → somewhat applicable; - → not applicable.

		Criteria				
Measure	**Source**	**Multi-item**	**Interview-coherent**	**All products**	**Similar context**	**Publication ranking**
PQ	Konuk, 2018; Stolz & Bautista, 2015	x	x	x	x	x
PI	Das et al., 2020	x	o	x	x	x
PEF	Magnier et al., 2016; Ye et al., 2022	x	x	x	x	x
Nat	Pula et al., 2014	x	x	o	o	-
Soc	Anagnostou et al., 2015	x	x	o	x	o
Ani	Lindeman & Väänänen, 2000; Stremmel et al., 2024	x	o	x	x	o
Cli	Stremmel et al., 2024	x	x	x	x	x
Res	Stremmel et al., 2024	x	x	x	x	x
Dur	Hazen et al., 2017	x	o	-	x	x
Rep	Hazen et al., 2017	x	o	-	x	x
Rec	*Self-developed*	x	x	o	-	-
Wglow	Das et al., 2020	x	x	x	x	x
Exper	Newell & Goldsmith, 2001	x	x	x	-	x
Indic	*Self-developed*	x	x	x	-	-
Innov	Brockman & Morgan, 2003	x	x	x	-	x
Effort	Bechwati & Xia, 2003	x	o	x	o	x
Stren	Mai et al., 2019	x	o	x	x	x

Some of the measures taken from literature were slightly adjusted in their item wording to better fit the products under investigation and to streamline the scaling across measures used: Some measures were originally applied to assess a specific product, for example the measure for Nat (Pula et al., 2014) was developed in the context of food products. To make it applicable to other products, especially the Sweater, item wordings were adjusted, i.e., added the word "materials" in addition to "ingredients", and used the word "substances" instead of "preservatives". Additionally, some measures originally used a 7-point semantic differential scale, meaning a scale where the endpoints are not "fully disagree – fully agree" but rather opposing adjectives, e.g., "socially responsible – socially irresponsible" (from the social friendliness measure adopted from Magnier et al. (2016)). To keep the answer scales consistent throughout the questionnaire, all bipolar statements were adjusted using only one adjective from the original measure and asking participants for their level of agreement. For all such statements, the "positive" endpoint of the original scale was taken (e.g., "fair", "beneficial"; instead of "unfair", "harmful"). Answer scales were kept consistent and only one adjective of the bipolar scales was chosen to limit the effort required from participants for filling in the questionnaire. Two of the measures taken from literature (Effort and Exper) contained reverse items, i.e., items that are framed in the opposite direction. The reverse items were negated statements, e.g., "(…) does not have much experience (…)" (Newell & Goldsmith, 2001). These were not adjusted but kept for consistency with the original measure. Table 17 shows the resulting list of items used in this thesis (German version in Appendix B).

Table 17: Measures and items used in the questionnaire.
Note: (*) = Reverse item.

Measure (source)	Items
PQ (Konuk, 2018; Stolz & Bautista, 2015)	1. This product is adapted to my needs. 2. This product is of high quality. 3. This product is a high standard product. 4. This product is of very good quality.
PI (Das et al., 2020)	1. I am likely to buy this product. 2. I am willing to buy this product. 3. I am inclined to buy this product.
PEF (Magnier et al., 2016; Ye et al., 2022)	1. This product is an eco-friendly product. 2. This product is a good example of an eco-friendly product. 3. This product is sustainable.
Nat (Pula et al., 2014)	1. It contains natural ingredients / materials. 2. It has undergone minimal processing. 3. It is free of chemical substances. 4. It contains no additives. 5. It is free of residues from chemical sprays and pesticides.
Soc (Anagnostou et al., 2015)	1. It has been created under fair working conditions. 2. It is a socially responsible product. 3. It is more beneficial to society's welfare than other products. 4. It has been traded under fair conditions.
Ani (Lindeman & Väänänen, 2000; Stremmel et al., 2024)	1. It has been produced in a way that animals have not experienced pain. 2. It has been produced in a way that animals' rights have been respected. 3. It has been produced considering animal welfare.
Cli (Stremmel et al., 2024)	1. It is beneficial for the climate compared to similar products. 2. It is good for the climate compared to similar products. 3. It contributes to climate protection. 4. It causes less greenhouse gases (e.g., CO2) than similar products
Res (Stremmel et al., 2024)	1. It is water-saving. 2. It is associated with low-impact land use. 3. It is resource-efficient.
Dur (Hazen et al., 2017)	1. Failure of this product does not occur often. 2. It has an adequate lifespan. 3. It is reliable. 4. It is a durable product.
Rep (Hazen et al., 2017)	1. Getting it repaired should not be an issue. 2. When it needs repair, it will not take long. 3. I can receive responsive repair services for this product. 4. I can easily find repair parts and services for this product.
Rec *Self-developed*	1. It can be reused. 2. It can easily be recycled. 3. It does not create unnecessary waste at the end of its useful life.
Wglow (Das et al., 2020)	1. I will feel good by purchasing this product because I will be able to help in protecting the environment.

Measure (source)	Items
	2. With the purchase of this product, I will have the feeling of contributing to the wellbeing of humanity and nature. 3. With the purchase of this product, I will feel better because it does not harm the environment.
Exper (Newell & Goldsmith, 2001)	1. Its producer has a great amount of experience with eco-friendly products. 2. Its producer is skilled with eco-friendly products. 3. Its producer has great expertise with eco-friendly products. 4. Its producer does not have much experience with eco-friendly products. (*)
Indic *Self-developed*	1. Eco-friendliness and quality have a lot in common. 2. My perceptions of eco-friendliness and quality are influenced by the same things (e.g., healthiness, durability, etc.). 3. Eco-friendliness and quality are related.
Innov (Brockman & Morgan, 2003)	1. This product challenges existing ideas for this product category. 2. This product offers new ideas to this product category. 3. This product is creative.
Effort (Bechwati & Xia, 2003)	1. Producers put more effort into creating eco-friendly products than conventional products. 2. Producers do not exert more effort in creating eco-friendly products than conventional products. (*) 3. Producers work harder to create eco-friendly products than conventional products.
Stren (Mai et al., 2019)	1. Conventional products are often more effective than eco-friendly produced products. 2. Eco-friendly products are less strong. 3. Regular products often are stronger than the eco-friendly alternatives. 4. Eco-friendly products are less durable, less robust, less tasty, etc. 5. Conventional products are often more durable, robust, etc., than products that are manufactured eco-friendly.

In addition to the construct measures described above, three additional items were included to identify participants who are not paying attention to the questionnaire. For that purpose, three attention check questions were included in the questionnaire, which is recommended for self-administered surveys (Berinsky et al., 2014) and applied in many studies (e.g., Atkinson & Rosenthal, 2014; J. C. Nunes et al., 2021). Two attention check questions were a *directed query* (Abbey & Meloy, 2017), where participants were asked to select a specific answer option (e.g., "disagree"). Thus, participants who fail to select the indicated options have most likely not read the statement and could overall distort the results due to their inattentiveness. The remaining attention check included in this questionnaire asked the participants for which product they had read the description, which can be classified as *manipulation check* (Abbey & Meloy, 2017). Participants who failed at least one attention check were excluded from analysis (see section 5.4.1 for more details).

5.2.2.4. Control variables

This section describes additional variables (mostly demographic) included in the questionnaire that might affect the correlations within the model, while there are no specific hypotheses about their impact. These variables are called control variables in this thesis, for which the information was collected at the end of the questionnaire. As shown in the literature review in section 3.2.3.1., previous research assessing the relations between PEF and PQ has not been fully conclusive about the effects of demographic variables like age, gender, or income. In addition, familiarity with the product, number and age of kids, country of longest residence, household size, living area (urban versus rural) and highest education are included as tentative control variables as these topics came up during the interviews or are commonly used in literature. For this thesis, a rather high number of control variables was chosen to be able to shed light on several topics and gain overall more clarity, while ruling out potentially confounding effects. The suggested format of each respective question for control variables was chosen to make answering as comfortable as possible for participants (e.g., income in categories rather than specific numbers). Table 18 shows an overview of the control variables included in the present study.

Table 18: Overview of control variables included (with items and answer options).

Measure (*type*)	**Items**	**Answer options**
Familiar (*5-point Likert-type scale*)	How familiar are you with products like the described one?	1: Not at all familiar – 5: Extremely familiar
Gender *(single choice)*	What is your gender?	• Female • Male • Non-binary
Age *(open text)*	Please indicate your current age in years.	*[Open text field]*
Child(ren) age *(multiple choice)*	Please indicate if you have children / how old your child(ren) is or are. Please choose all options that apply.	• I do not have a child • < 5 years • 5-9 years • 10-14 years • 15-19 years • >19 years
Household size *(open text)*	Please indicate how many people live in your household.	*[Open text field]*
Living area *(single choice)*	Which option describes your living area best?	• Urban • Rural
Highest education *(single choice)*	What is your highest level of education?	• Secondary school • Vocational / technical / trade training • Completed university degree (e.g. Bachelor, Master, Diploma) • Completed doctorate degree
Residence *(single choice)*	In which country have you lived for most of your life? If "other", please specify.	*[List of European countries with highest population, and option to specify "other"]*
Household income *(single choice)*	What was your household net income last calendar year in €?	• < 25000 • 25000 – 49999 • 50000 – 74999 • 75000 – 99999 • > 99999 • Prefer not to disclose

Beyond including control variables like the ones above, some authors include a measure for social desirability. *Social desirability* "is the general tendency of individuals to present themselves in a manner that makes them look positive with regard to culturally accepted standards of behavior" (Chung & Monroe, 2003, p. 292). Thus, social desirability can impact how participants answer survey questions by trying to present themselves in an eco-friendlier way. A measure to control for social desirability has not been included in the survey, because of two key reasons: 1) Due to the web-based, self-administered and anonymous format, social

desirability is less likely to occur than e.g., in a laboratory setting or during personal interviews (Kreuter et al., 2008; Tourangeau et al., 2013); 2) Established measures of social desirability consist of at least 6 items (Fischer & Fick, 1993; Reynolds, 1982). Including those would increase the length of the questionnaire considerably and thus potentially reduce the answer quality or increase drop-out rate as the number of overall questions can influence how many people drop-out (Galesic, 2006).

5.2.3. Data analysis

Data analysis occurred in four stages: First, data was prepared (e.g., re-coding of reverse items), overall data quality and suitability for further analyses was assessed and adjustments made as needed. Second, factor analysis was conducted to confirm the underlying structures and multi-item measures. Third, regression analyses were conducted to prepare the SEM analyses. Fourth and finally, structural equation models were assessed using the PLS SEM method. Analyses were conducted in R using version 4.4.2 and R studio version 2024.12.0.467. In case of issues related to specific code segments or functionalities in R, ChatGPT (OpenAI, 2024) was selectively leveraged to support issue resolution (e.g., by pointing out potential mistakes in individual code segments).

The following sections introduce factor analysis (see section 5.2.3.1), regression (see section 5.2.3.2) and SEM (see section 5.2.3.3) in more depth before presenting the results of all analysis stages subsequently (starting at section 5.3).

5.2.3.1. Factor analysis

Factor analysis is a prominent method used to assess underlying data structures. It has been extensively applied in research across various fields for many decades. Broadly, factor analysis is categorized into two main types: *exploratory factor analysis* [EFA] and *confirmatory factor analysis* [CFA], each with different subcategories and applications. Both EFA and CFA are used to understand factor structures between observed indicators and latent variables.

EFA is typically used to explore potential factor structures and conduct data reduction without prior hypotheses by identifying latent variables from observed indicators; while CFA is applied when there is a predefined model, i.e., a hypothesized factor structure based on theoretical considerations (Bühner, 2011). For the research at hand, CFA was chosen because there is a theoretical model and factor structure, and additionally, most measurement scales for the latent variables have been taken from established literature so that the relations between

items and latent variables for these measures have been previously validated. Additionally, after confirming the proposed constructs, CFA often leads to SEM (Hair et al., 2020), as is intended in this thesis.

Before conducting the CFA, suitability of the data needs to be assessed. This is typically done via the *Kaiser–Meyer–Olkin* [KMO] test of sampling adequacy (Kaiser, 1970) and Bartlett's test of sphericity (Bartlett, 1950). The KMO test "yields an assessment of whether the variables belong together psychometrically and thus whether the correlation matrix is appropriate for factor analysis" (Dziuban & Shirkey, 1974, p. 359). According to Dziuban and Shirkey (1974), Kaiser (1974) defined values in the 0.90s as marvelous, in the 0.80s as meritorious, in the 0.70s as middling, in the 0.60s as mediocre and below 0.5 as unacceptable. For the study at hand, the KMO measure was in a marvelous range for the overall data (KMO = 0.93; and individual scores for indicators all ≥ 0.78) as well as the product specific data (KMO Chocolate = 0.91; KMO Sweater = 0.90). Bartlett's test of sphericity confirms whether the variables of interest are sufficiently correlated to indicate a certain level of redundancy, which is needed so that reducing the dimensionality via factor analysis is possible (Eid et al., 2011). The test was significant for the overall data as well as the individual product data indicating a good fit for factor analysis. Additionally, factor analyses typically assume multivariate normality of the indicators. The Mardia test was conducted as is commonly recommended (Bühner, 2011). This test revealed that multivariate normality was not given. To address non-normality, a correction method can be applied to the *maximum likelihood* [ML] estimation method. ML estimation is commonly used in CFA and is robust to certain violations of normality, especially when combined with appropriate corrections (Bagozzi, 2010). As correction method, the Satorra-Bentler correction (Satorra & Bentler, 1994) was chosen in the present work. This method is considered the most widely used correction applied to provide more accurate results under conditions of non-normality (Du & Bentler, 2022; Jobst et al., 2022).

Thus, after having established suitability of the data for factor analysis as described in the last section, a CFA was conducted using the lavaan package (Rosseel, 2012) in R. The objective of conducting factor analysis in the current research is twofold: First, conducting a factor analysis should investigate the underlying factor structure or, in other words, confirm that all indicators that intend to measure a specific construct actually measure that construct. This can be done via assessing factor loadings and cross-loading with other constructs. Based on the results of the factor analysis, specific items might be excluded from further analysis.

Second, the hypothesized dimensionality of PEF and specific dimensions per product should be inspected.

To evaluate the results from CFA, several goodness of fit measures exist and will be applied. A common set of goodness of fit measures evaluate the results on overall model level. Specifically, they evaluate the overall consistency of the model with the data. Indicator level measures of fit evaluate individual factors and indicators. As local measures, factor loadings and *modification indices* [MI] will be used. Modification indices "are single-degree of freedom χ^2 tests that show what would happen to the overall global χ^2 if an additional arrow would be added to the model" (Thoemmes et al., 2018, p. 2). A high value in a modification index indicates that local fit can be improved. Specifically, the value of a modification index tells the user how much the overall χ^2 of the complete model would approximately change by freely estimating for example a specific covariation between indicators or additional loading of an indicator (Brown, 2023).

In addition, the in section 5.2.2.1 explained criteria for internal consistency reliability and convergent validity will be applied in this study. Since cross-loadings with other factors (discriminant validity) are not directly available as results of CFA, this will be addressed by leveraging MI. These MI allow for example to identify where a cross-loading of an indicator with one or more other factors would improve the overall model substantially. See Table 19 below for a summary of measures of fit. Thresholds for the model-level measures are based on Bagozzi (2010). Regarding the threshold for factor loadings, there is no clear guidance, but values of 0.3 or 0.4 are often considered the minimal acceptable boundary (e.g., Heene et al., 2011) and 0.4 is described as commonly applied threshold (Deo & Prasad, 2024). Since MI are also used to identify indicator-level misspecifications, 0.4 was chosen as minimal acceptable level for this thesis.

Because χ^2 is proportional to the sample size and is significant when the sample size is large even though the model is fitting well, Bagozzi (2010) recommends in such cases to rely only on the other fit indices. In that case, ideally RMSEA, SMSR, CFI and TLI all are in their respective recommended ranges: *Root mean squared error of approximation* [RMSEA] considers the χ^2, the *degrees of freedom* [df] and the sample size. *Standardized root mean square residual* [SMSR] provides a standardized measure for effect size of misfit, while RMSEA is an unstandardized measure (D. Shi et al., 2018). *Comparative fit index* [CFI] is normed (within the range of 0-1) and compares the empirical model against a basis model while

accounting for degrees of freedom (L. Hu & Bentler, 1999). Similarly, the *Tucker Lewis index* [TLI] (Tucker & Lewis, 1973), also known as non-normed fit index, compares empirical and basis model, but it is not adjusted to the complexity of the model (L. Hu & Bentler, 1999), thus often favoring simpler models. It is relatively independent of sample size (Marsh et al., 1988). As compared to the other measures of fit, *Akaike information criterion* [AIC] and *Bayesian information criterion* [BIC] are relative measures that have not much meaning in isolation but can be used to identify the best fitting model out of several alternatives (Schwarz, 1978). Within this thesis, AIC and BIC are used to compare if specific modifications to a CFA model improve the overall model.

Table 19: Measures of fit for CFA applied in this thesis.
Sources: Based on Bagozzi (2010) and (Deo & Prasad, 2024).

Measure	Threshold
Model-level	
χ^2	n/a (should be non-significant, $p \geq 0.05$)
RMSEA	≤ 0.06
SMSR	≤ 0.08 (better ≤ 0.07)
CFI	≥ 0.95 (or ≥ 0.90)
TLI	≥ 0.95 (or ≥ 0.90)
AIC / BIC	n/a (lower values are better)
Indicator-level	
Factor loading	≥ 0.4 (needs to be significant, $p < 0.05$)
Modification indices	n/a (lower values are better)

After having conducted CFA to assess and improve (where needed) the internal consistency, reliability and validity of the multi-item measures, the individual indicators will be combined into a mean scale for each variable. These mean scales are needed for further application in regression analysis. Additionally, the newly created scales are also created in a standardized version. This procedure is often recommended to increase interpretability and meaning of data analysis results (e.g., Bühner & Ziegler, 2009; Eid et al., 2011; Hair et al., 2021). Standardization (also called z transformation) of an observed variable is conducted by subtracting the mean from each individual observation and dividing the result by the standard deviation of that item (Eid et al., 2011). Standardized variables have a mean of 0 and standard deviation of 1 making results from e.g., regression or SEM analyses are easier to interpret. Let's take a path coefficient of 0.4 for a linear model with one predictor and one dependent variable as example: It can be interpreted that the dependent variable increases by 0.4 (standard

deviation units) if the predictor increases by 1 (standard deviation units). Note that the original shape of the distribution is not affected by standardizing. There are alternative ways of transforming variables that also change the shape of the distribution, e.g., logarithmic, and can be used to achieve normality of originally non-normal distributed data. However, logarithmic transformations are much more useful for reaching normality if data are positively skewed as opposed to negative like observed in the current data. A more flexible way of transforming data is via so called Box-Cox transformations (Box & Cox, 1964) because it allows to first identify the best parameter for reaching normality and then applying a transformation accordingly. Since any transformations of data have implications on the interpretation, no transformation beyond standardizing is conducted in this work. In the current work, standardization was applied separately to each product dataset and only after, the datasets were joined together to create the complete standardized dataset. This approach was chosen so that further analyses and interpretations could be conducted in a meaningful way for each product. Because the mean and standard deviation of indicators differed between products, standardizing on the combined dataset could distort the individual product data and thus impact results. The results from factor analysis as described here are presented in section 5.3.3. After the factor analysis, regression analyses are conducted as described in the next section.

5.2.3.2. Regression

This section introduces linear and multiple linear regression and how it is applied in this thesis. Regression analyses are conducted in this thesis to identify which control variables should be included in the SEM analyses, if there are any influential observations that should be excluded for the SEM analyses and if there are any hypothesized effects that should be excluded from the SEM analyses to reduce the overall model complexity. Beyond investigating the hypotheses, regression analyses are intended to provide a first assessment of the effects of a *multidimensional measure of PEF* [PEF_dim] in comparison to a *unidimensional PEF measure* [PEF_uni].

Regression analysis is a commonly applied methodology of data analysis in business, marketing, psychology, and many other research fields as well as in practical applications. Regression analysis allows to model and interpret various correlations between one or multiple independent variables / predictors and a dependent variable. It also allows to investigate moderation and mediation. A key requirement for linear regression to be applicable is a linear relationship between the independent and the dependent variables. In order to test whether linearity between an independent and the dependent variable can be assumed, visual inspection

via residual versus fitted plots and smoothed scatterplots was conducted. Visual analyses indicated that bivariate relations between the constructs can be considered linear even though there seemed to be some deviations in areas with very few data (e.g., values below 3 for PQ). In the following, linear regression will be explained in more detail.

A *standard linear regression* in its simplest form with only one predictor variable can be described by the following formula:

$$Y = \beta_0 + \beta_i x_i + \varepsilon$$

for which in this thesis Y is PQ (the dependent variable), β_0 is a constant term, x_i is PEF (the independent variable / predictor), β_i is an (unknown) regression weight and ε (error term) are mutually independent random variables which account for the variation in Y not explained by the independent variable. To assess the impact of the predictor on the dependent variable, the regression weight needs to be estimated. Typically, β_i is chosen so that the model fits the data best as operationalized of minimizing the residual sum of squares. Standard linear regression can be extended to *multiple linear regression* [MLR] by including multiple variables (i.e., additional x and respective β) to understand more complex data structures more comprehensively (Bühner & Ziegler, 2009; Montgomery et al., 2021). MLR is very commonly applied in marketing research to analyze the relationships between a dependent variable and multiple independent variables and allows to assess the influence of each predictor on the dependent variable while controlling for the influences of the other variables included in the regression. Beyond the main effects of multiple predictors, MLR can be used to analyze the interaction effects of predictors on the dependent variable (in this case PQ) depending on the level of another predictor. Interaction terms thus allow to investigate moderating effects within a model.

For MLR analysis, the data should fulfill certain criteria to ensure that results are valid. These criteria include homoscedasticity (the constant variance of errors), independence and normality of error terms, and absence of multicollinearity, in addition to the linearity (Eid et al., 2011). Within this thesis, those criteria will be assessed for every regression model. As indicated in the beginning of this study, linearity is visually assessed and where needed further inspected via quadratic regression terms. Homoscedasticity is assessed using the *Breusch-Pagan test* [BP] (Breusch & Pagan, 1979). Independence of errors is assessed via the *Durbin-Watson test* [DW] (Durbin & Watson, 1950, 1951, 1971) which is considered a powerful test and robust against non-normality (Bartels & Goodhew, 1981). For normality of errors, *quantile-*

quantile plots [Q-Q plots] of the residuals are visually inspected and the *Shapiro-Wilk test* [W] is used on residuals (Shapiro et al., 1968). Multicollinearity results from highly correlated independent variables and is assessed via the *variance inflation factor* [VIF] of each construct in the regression (or SEM) model. Values above 10 are typically considered problematic (Bühner & Ziegler, 2009; O'brien, 2007). The results of these analyses are described together with the overall regression results in section 5.4.4. Heteroscedasticity and non-normality of errors are commonly observed in regression analyses (Hair et al., 2013), so that potential remedies are considered already at this stage. According to Eid et al. (2011) non-normality of errors is less worrisome for rather large sample sizes so that slight deviations from normality do not require adjustments given the sample sizes available. However, to ensure quality of the results, bootstrapping of the regression results to obtain distribution independent results will be conducted if visual inspection of Q-Q plots of the residuals strongly suggest non-normality. Bootstrapping estimates the coefficients and standard errors by repeatedly resampling data with replacement from the original data, which allows to get robust and reliable estimates (Eid et al., 2011). Bootstrapping as non-parametric method provides a *confidence interval* [CI] for the coefficient: If 0 is not within the CI, then an effect is considered statistically significant. If the results are rather normally distributed but show heteroscedasticity, a robust standard error analysis will be applied to the MLR. Since the sample size in the present study exceeds 250 in all scenarios, White's estimator (H. White, 1980) is chosen to create the heteroscedasticity-consistent covariance matrix (Long & Ervin, 2000) and thus the robust standard errors.

In combination with assessing the statistical criteria or assumptions (e.g., homoscedasticity) as described above, the regression results are investigated in terms of outliers and *influential observations* [InfObs]. Outliers in a regression model can be identified via a studentized residual with a value < -3 or > 3 (Eid et al., 2011). InfObs describe any observations (i.e., participants) that have a strong impact on the estimated regression coefficients / results (Eid et al., 2011). Potentially InfObs can be identified via different methods (Eid et al., 2011; Hair et al., 2018): A) The hat value, also called leverage, of an observation indicates the distance from the mean center of the independent variables of all other observations. It can also indicate that an observation has a comparingly high impact on the estimated value of the dependent variable; B) DFBETA and DFFIT values indicate the difference in regression coefficients (DFBETA) and the difference in the estimated value of the dependent variable (DFFIT) between running a regression with the respective observation or running a regression without considering that observation – both DFBETA and DFFIT can be assessed unstandardized or

standardized (SDFBETA and SDFFIT); C) Cook's distance, being influenced by the changes in the dependent variables when excluding an observation and the leverage of that observation; and D) COVRATIO considers all coefficients collectively and estimates the impact of an observation on the standard errors of the coefficients. According to Hair et al. (2018), Cook's distance is considered the most representative to assess the influence of an observation on the overall model fit. Table 20 summarizes the common measures and their thresholds as suggested by Hair et al. (2018) and Eid et al. (2011). However, the authors also state that there are no clear rules and it remains subject to the interpretation of the researcher – always aiming to increase validity and generalizability by any modifications.

Table 20: Measures and thresholds to identify potential InfObs.
Sources: Based on Hair et al. (2018) and Eid et al. (2011).
Notes: N = sample size; i = number of independent variables.

Measure	Suggested threshold
Leverage	2i/N or 3i/N
SDFBETA	$\pm 2/\sqrt{N}$
SDFFIT	$2\sqrt{((i+1)/(N-i-1))}$
Cook's distance	$4/(N-i-1)$
COVRATIO	$1 \pm 3i/N$

5.2.3.3. Structural equation modelling

This section describes another method applied within this thesis, SEM, which "has become a quasi-standard with respect to analyzing cause-effect relationships between latent variables" (Hair et al., 2017, p. 616). Simply put, SEM combines CFA and MLR, allowing to simultaneously investigate multiple relationships between latent variables by integrating measurement validation with path analysis (Eid et al., 2011). In SEM, a distinction between exogenous and endogenous variables can be made: Exogenous variables explain other variables in the model (e.g., Y1 and Y2 in Figure 13) while endogenous variables are explained by other variables in the model (e.g., Y2 and Y3 in Figure 13) (Hair et al., 2022). Beyond that, every SEM consists of measurement and structural models. A structural model defines the relationships between latent variables while a measurement model defines the relationships between observed variables and a latent construct (Hair et al., 2022). Both structural and measurement models can contain error terms which stand for the unexplained variance of estimated models (Hair et al., 2022). In the structural model, endogenous variables include error terms (labelled "z" in Figure 13). In the measurement model, indicators for reflective constructs

include error terms (labelled "e" in Figure 13). Note that constructs measured by a single item (e.g., Y3) or formatively defined constructs do not require a measurement error term.

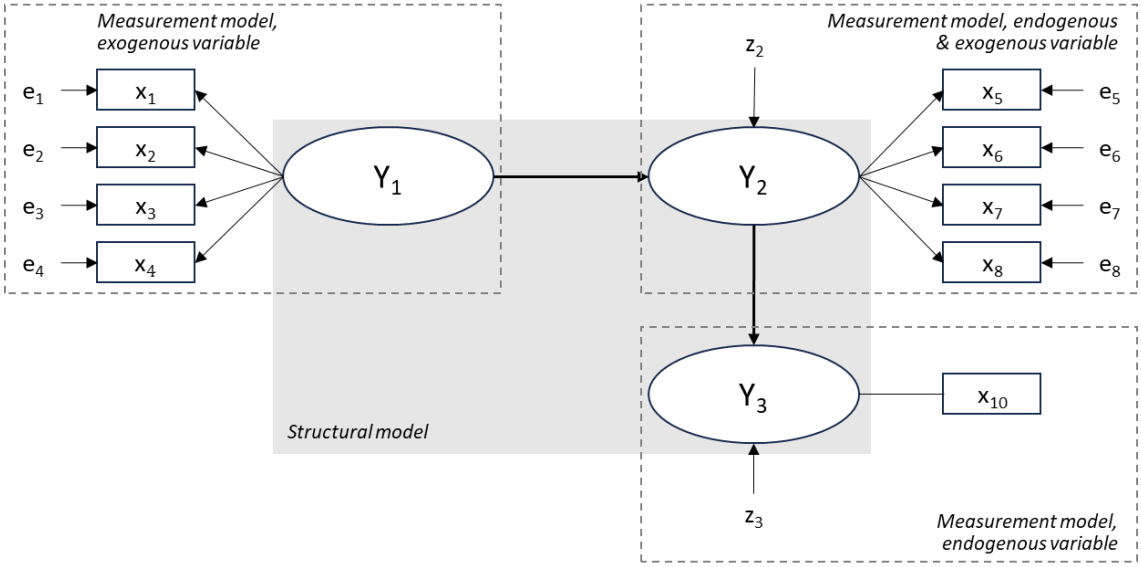


Figure 13: Visualization of exogenous and endogenous reflective latent variables, and exemplary measurement and structure models.
Source: Adapted from Hair et al. (2022).

Two main approaches of SEM are available to researchers: Factor or *covariance-based SEM* [CB SEM] and PLS SEM. Previously, CB SEM used to be more widely applied in marketing research but PLS SEM is becoming increasingly popular (Hair et al., 2017). PLS SEM uses ordinary least square regression aiming to minimize the residual variance of the endogenous constructs, which means maximizing the *explained variance* [R^2] of the endogenous constructs (Hair et al., 2021). Between the two approaches, PLS SEM is in general more flexible, e.g., it does not assume that a common factor explains all covariation in the indicator sets. Specific properties described by Hair et al. (2021, 2022) led to choosing PLS SEM over CB SEM for the current study, including:

- It is applicable to non-normal data (no distributional assumption)
- It works well with complex models and relatively smaller sample sizes
- It handles constructs measured with single items, as well as reflective and formative multi-item measures
- It facilitates specification of (multiple) interaction effects
- It can be used with nonlinear effects (to a certain extent)

- It is typically favored in a somewhat explorative setting rather than confirming long established theories - which applies to the current situation as the GT model contains several new constructs not previously assessed in the context of PEF and PQ

Overall, even though CB SEM and PLS SEM differ statistically, PLS SEM estimates can approximate CB SEM results well, especially when normality is not given, model complexity is high or sample size requirements for CB SEM are not given (Hair et al., 2011). Comparing PLS SEM to MLR, there are several advantages: For example, PLS SEM can effectively account for measurement errors via weighted composites whereas MLR uses sum scores with equal weights for all indicators (Hair et al., 2021). However, PLS SEM also comes with some potential disadvantages compared to CB SEM. Most importantly, established global measures to evaluate goodness of model fit (like RMSEA, CFI, etc.) cannot be applied to PLS SEM (Hair et al., 2022).

In the current study, bootstrapping is applied to the model and resulting CIs are used to determine if a specific path is statistically significant. Bootstrapping is a common method which repeatedly resamples data with replacement from the original data and thus allows to get robust and reliable estimates (Eid et al., 2011). In line with recommendations by Hair et al. (2021), bootstrapping is conducted with 10,000 iterations for SEM related analyses in this thesis. In addition to bootstrapping, further criteria are used to assess the suitability of the model, namely the VIF to assess multicollinearity as well as reliability and validity measures described in section 5.2.2.1 (namely Cronbach alpha, rhoC, AVE, FLC, HTMT).

To conduct a SEM analysis, the measurement model and the structural model need to be defined: The measurement model defines how the latent constructs are measured (e.g., reflective / formative and which indicators are used per construct); the structural model defines the relationships / paths between the latent constructs. In addition, the weighting scheme needs to be specified. Within this thesis, the path weighting scheme recommended by Hair et al. (2021) is used which maximizes the explained variance of endogenous constructs. The measurement and structural models under investigation within the current research require some extensions of PLS SEM beyond including linear main effects. Specifically, interaction effects and hierarchical constructs (also called second-order constructs, in this case PEF as measured by its dimensions) need to be modelled. In the next paragraphs, interaction effects and nonlinear effects are described in the context of PLS SEM.

Interaction effects. Like in regression analyses, interaction terms can be used to assess moderation. However, different to regression analyses, the constructs are measured in SEM via individual items instead of using predefined sum scales. Thus, there are different methods of implementing interaction terms in SEM: product indicator, two-stage and orthogonal. For the purpose of this work, the orthogonalizing method is chosen, because 1) the model contains only reflective constructs and 2) this methods reduces the risk of collinearity issues in interpretation due to the interaction term by setting the interaction term uncorrelated with the predictor and moderator (Becker et al., 2018). According to Hair et al. (2021), including interaction effects can change how a main effect of the predictor should be analyzed and they recommend first assessing main effects in a model without interactions and only afterwards extending the model. Thus, first only the main effects will be interpreted and then interaction terms added. In terms of evaluating the relevance of interaction effects, the effect sizes should be assessed in addition to the significance level. Effect size indicates the change in explained variance of an endogenous construct when a specific path (in this case moderating effect) is included or excluded (Hair et al., 2021). In moderation analysis, values of 0.005, 0.01, and 0.025, respectively, should be considered as small, medium, and large effect sizes (Hair et al., 2021).

Hierarchical constructs. Hierarchical constructs are also called higher-order or second-order constructs and consist of "two layered structures of constructs" (Hair et al., 2024, p. 33). Within this thesis, PEF as multidimensional construct (also referred to as *PEF_dim*) is represented via several dimensions identified in the interviews (e.g., Nat or Soc). Thus, it is defined as reflective-reflective hierarchical construct, meaning that the lower order constructs (e.g., Nat) are measured reflectively and the higher order construct as well. According to Hair et al. (2024), there are four approaches to establish the measurement model of the higher-order construct: Repeated indicators, extended repeated indicators, embedded two-stage and disjoint two-stage. Within both versions of the repeated indicators approaches, the higher-order construct is represented by each individual indicator of the relevant lower-order constructs, meaning that these indicators are used twice. For the two-stage approaches, in a first stage construct scores for the lower-order constructs are estimated and then these are used as indicators for the higher-order construct in the second stage. In line with recommendations from Hair et al. (2024), the disjoint two stage approach was chosen and incorporated in the following way: First, the latent variables scores of the applicable dimensions are calculated via PLS SEM containing all other constructs (except for the hypothesized interaction terms and any potential socio-demographic control variables). These scores are then added as new variables to the

dataset. In the second step, the hierarchical construct (i.e., PEF_dim), is defined using those scores and treating them as observed variables and thus as indicators. In this step, all other constructs and hypothesized paths are included as well.

In addition to the so far described set-up and assessment of PLS SEM analyses, these models can be analyzed regarding their predictive power, i.e., how well the model can predict dependent variables for new datasets. In this thesis, the predictive power of the models is assessed to confirm the overall relevance of the models. This is implemented via 10-fold cross-validation as recommended by Shmueli et al. (2019), meaning that the full dataset is split into 10 equal subsamples of which nine are used for training the model and the remaining subsample is used as test dataset for prediction. To avoid extreme partitioning and increase precision, the 10-fold cross-validation is repeated 10 times (Witten et al., 2016) in this thesis and the averages are considered for evaluation. The precision of the prediction and thus the predictive value of the model is then assessed by comparing the prediction error of the PLS SEM with a benchmark based on a linear regression model (Shmueli et al., 2019). The most common prediction error used for this is the *root-mean-square error* [RMSE] according to Hair et al. (2021), which is used in this thesis as main indicator and complemented by the *mean absolute error* [MAE]. Since the model under investigation contains mediating effects, it needs to be decided how the mediators are considered for predictions: In line with recommendations by Hair et al. (2021), the "direct antecedents" approach is chosen. It is recommended to focus the assessment of prediction error only on the most relevant endogenous (i.e., dependent) constructs and not on all endogenous construct (Hair et al., 2021), so that only PQ and PI will be assessed in this study. The predictive power of the model is considered high, if all the PLS SEM prediction errors of all indicators of the key endogenous construct(s) are lower than their linear regression benchmark errors (Shmueli et al., 2019). Similarly, predictive power is considered medium if most or the same number of indicators have a lower PLS SEM than linear regression prediction error. If less than half of the indicators of the key dependent construct(s) have lower PLS SEM than linear regression prediction error, the overall predictive power is considered low.

After having introduced the statistical methodologies applied in the current section, the next section describes the data collection and preparation, before the results from PLS SEM modelling are presented in section 5.4, followed by a discussion of the results in section 5.5 and a summary of quantitative research in section 5.6.

5.3. Data collection and preparation

In this sub-chapter, first the results for the overall data will be reported before specific results for each product separately are presented. The total dataset was created by combining the results from all sources (as described in section 5.4.1). Datasets for individual products (Chocolate and Sweater) were created by splitting the total dataset after initial (e.g., descriptive) analyses and corrections (e.g., excluding responses with failed attention check, handling missing values) had been made.

5.3.1. Data collection and sample

Within this section, the distribution of the questionnaire, inclusion / exclusion of participants and the resulting sample are described. Data were collected via an online questionnaire, leveraging Unipark software (2024) as survey platform. Data collection occurred between October and December 2024. The questionnaire was fully anonymous. Participants first read information about the purpose, scope and procedure of the study and the handling of data, and then consented to participate (including data usage). Participants were informed that they can drop out at any time without facing consequences and were provided with a contact to reach out to in case of questions. There were no upfront restrictions on eligibility of participants but rather a broad set of demographic control variables included at the end of the questionnaire to understand their potential influences (e.g., age, gender and household size).

Especially in online based surveys, invalid responses or more specifically meaningless responses from participants who are inattentive or unwilling to give valid responses often bias results and should thus be excluded from analyses (Leiner, 2019). Leiner (2019) investigated different strategies to identify meaningless responses and identified consideration of completion times and attention check questions as well performing. Thus, for this thesis, some submitted answers were excluded post hoc due to failed attention check questions or extreme response times. Response times were cut-off at four minutes, excluding everyone who completed the questionnaire faster. Four minutes was chosen as threshold because this was the value of the average duration minus one standard deviation (after excluding all answers with any failed attention check). Next, participants who took a break and then resumed the survey later were excluded if they filled the survey beyond a 60-minute interval. Taken together, four minutes was used as lower cut-off time and 60 minutes was used as upper cut-off time. In addition to excluding meaningless responses, any observations with missing values in the indicator items (see section 5.2.2.3) were excluded from further analyses. This decision was taken in line with general guidelines about handling of missing data (Hair et al., 2013).

However, some of the remaining observations still had missing values in one or more of the control variables, mostly in income (where the answer "prefer not to disclose" was considered as missing value). All control variables were analyzed for missing data and outliers / unexpected values. Where unrealistic values were found (e.g., household size of 0, or age of 570), the respective observations were individually inspected further and unrealistic values then defined as missing value (i.e., NA), assuming participants made unsystematic mistakes when filling these questions. Missing data were inspected for the whole dataset and were minimal since all questions were mandatory to answer, yet some questions could be skipped after a prompt. Finally, participants' comments at the end of the survey were reviewed and observations were excluded from analyses if participants indicated that they did not understand questions or wordings or had other difficulties, e.g., "The processes was a little confusing and a little redundant" or "It was difficult to decide which questions were not applicable in a few pages ago. for example should I have guessed the working conditions...".

Participants were recruited via three different channels: 1) The questionnaire link was distributed via the author's network, especially using social networks (e.g., LinkedIn, Facebook, Instagram, WhatsApp, nebenan.de); 2) In addition, online research platforms SurveyCircle (SurveyCircle, 2024) and SurveySwap (SurveySwap B.V., 2024) were used to recruit participants (unpaid); 3) Another part was recruited using the paid research platform Prolific (Prolific, 2024). The following paragraphs explain the sampling via the three different channels and then present the total resulting sample.

Personal network. Data collection via personal online and offline social network (including snowballing) was conducted in October and November 2024. 222 participants submitted the survey. The submission rate was ~67%, meaning ~67% of all people who read and gave their consent actually submitted the survey. Of all submissions, 26 (~12%) failed at least one attention check question, none contained a missing value in an indicator item and seven (~3%) additional responses were excluded due to their response times. In addition, seven participants (~3%) were excluded due to their comments at the end of the survey. Thus, a total of 182 answers remained after initial cleansing. Out of those, 94 (~52%) were based on the Chocolate scenario and 88 (~48%) on the Sweater scenario.

Unpaid online research platforms. Data collection in unpaid online research platforms occurred in November 2024. SurveyCircle and SurveySwap are both point-based systems, where participants earn points for filling surveys which in turn can be used to gain participants

for their own research. 314 participants submitted the survey, with a completion rate of ~94% of all people who read and gave their consent. Of all submissions, 74 (~24%) failed at least one attention check question, two (< 1%) had a missing value in an indicator item and ten (~3%) additional responses were excluded due to their response times. Additionally, one (< 1%) participant was excluded due to the comment provided. Thus, a total of 227 answers remained after initial cleansing, of which 113 (~50%) answered for the Chocolate, and 114 (~50%) answered for the Sweater.

Prolific. Data collection in this paid online research platform occurred in November 2024, in several waves during different days and times of the day. Participants on this platform get paid a certain pre-determined amount to fill in the questionnaire. For the study at hand participants received the equivalent of £1.90 (amount in British Pound due to the platform being based in the UK). 388 participants submitted the survey, with a completion rate of ~98% of all people who read and gave their consent. Of all submissions, 29 (~7%) failed at least one attention check question, eight (~2%) had missing values in any indicator item and 16 (~4%) additional responses were excluded due to their response times. Another six participants were excluded from further analyses based on the comments they added at the end of the survey. Thus, a total of 329 answers remained after initial cleansing, of which 163 (~50%) had the Chocolate as product and 166 (~50%) the Sweater.

Differences between recruitment channels. To analyze differences between the individual sub-samples as defined via their recruitment channels (i.e., personal network, unpaid or paid research platforms), several statistical tests were conducted. First, variables were assessed in terms of their normality and then homogeneity of variances as these are requirements for certain tests to compare groups (Eid et al., 2011). For normality assessment, several analyses were chosen which are described in more detail in section 6.3.5.1. In short, skewness and kurtosis were inspected as well as Q-Q plots before a formal test. A Q-Q plot shows the sample quantiles compared to quantiles as expected in case of normal distribution so that a straight line indicates normality (Oppong & Agbedra, 2016). As formal test, the Shapiro-Wilk test was chosen which is often considered the best performing test for normality (Patrício et al., 2017; Razali & Wah, 2011; Shapiro et al., 1968; Yap & Sim, 2011). While skewness and excess kurtosis for most indicators were between -2 and +2, indicated the Shapiro-Wilk test a lack of normality. In general, some divergence from normality has been regarded as not concerning for large sample sizes (Eid et al., 2011). Nevertheless, homogeneity of variances was assessed using Levene's test as this test does not require normality and can be applied to

multiple groups (Eid et al., 2011). Several variables showed heterogeneity of variances. Thus, differences between sub-samples (per channel) were analyzed using Kruskal-Wallis and Dunn's post hoc test for numeric variables. Kruskal-Wallis test is a non-parametric test for differences between multiple groups and typically preferable over other methods (like analysis of variance / ANOVA) if normality is not given (Eid et al., 2011; Hecke, 2012; McCrum-Gardner, 2008). Homogeneity of variances is preferred also for the Kruskal-Wallis test, however since there seemed to be no big differences in mean or median values (as visually inspected via boxplots), homogeneity of variances is not necessarily required for this test (Eid et al., 2011). Dunn's post hoc test is commonly applied after the Kruskal-Wallis test to identify which groups are differing in case of a significant Kruskal-Wallis test (e.g., Alexis Dinno, 2014; Gutiérrez et al., 2023). For categorical variables, chi-square and pairwise Fisher's exact post hoc test were conducted (Rufibach, 2009). Across variables and tests, respective Bonferroni corrections were applied to post hoc tests to account for multiple variable testing.

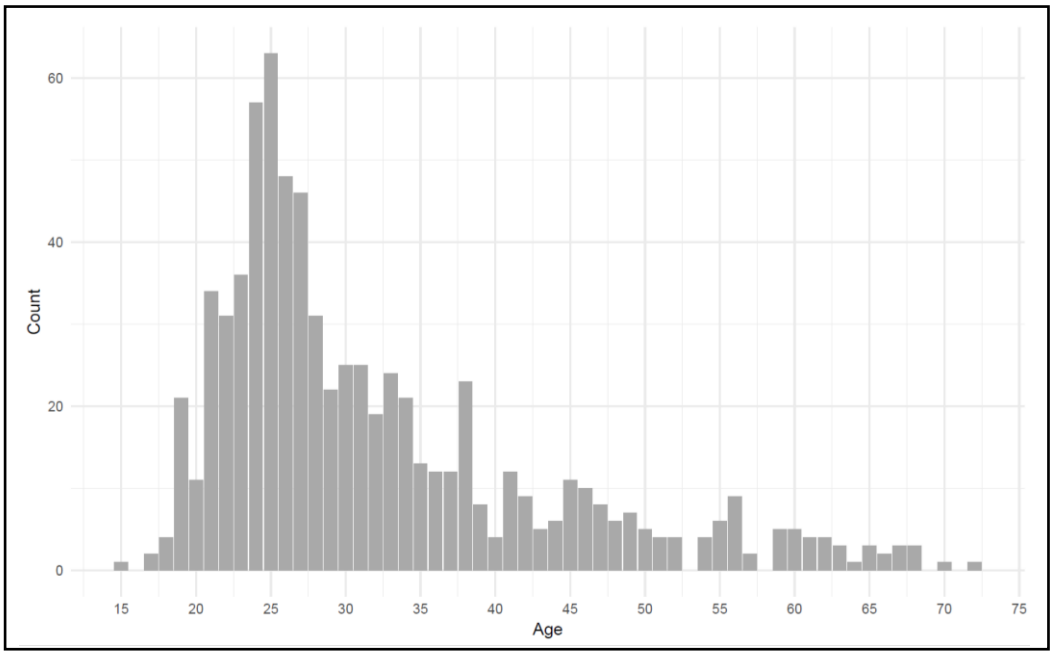
Results showed significant differences between the channels in demographic control variables and construct indicators. Demographic differences were for example found regarding gender, age, or income. These demographic differences in the subsamples were expected as the data collection from different channels likely represents different groups of society (e.g., unpaid research platforms tend to be used by students which has implications for e.g., age or income). Thus, combining different samples increases the diversity of the total sample which is helpful to reflect a broader population (note that the data in this study is not representative). Differences in key indicator variables (like indicators for PQ, Exper, etc.) were visually inspected via boxplots and outliers identified. Further inspection of the data did not provide specific reasons to exclude outliers at this stage from the subsequent data analyses. Differences between samples are reasoned to be related to the different demographic characteristics and overall sample composition between recruitment channels and show the value of combining different sub-samples for the following analyses. Thus, all valid responses were included in the subsequent data analyses.

Total sample. Based on the data collection and cleansing as described above, a total sample size of 738 observations is considered for data analysis. Out of these observations, 370 (~50%) are in the Chocolate condition and 368 (~50%) in the Sweater condition. A sample size of 311 had been determined as minimum sample size per product condition via a priori planning in G*Power (Faul et al., 2009) to identify small to medium effects with an α and β error of 5% each. Thus, the sample size requirement is fulfilled.

In a later stage of the analysis (after regression modelling to identify the relevant control variables, see section 5.3.4), six observations are excluded from the sample of 738 observations for the final SEM due to missing data in the relevant control variables. This leads to a final sample for hypotheses assessment of 732 observations (367 for Chocolate and 365 for Sweater. The descriptive analyses in the following are based on the full sample of 738 observations because these are used in the subsequent regression analyses and the differences compared to the dataset of 732 observations are minimal.

To start descriptive analyses, appropriate data types (e.g., factor, ordinal, etc.) are assigned for control variables and each control variable is visually inspected using histograms to gain insights into the sample composition. Mean age of participants is 32 years (median = 28 years, *standard deviation* [SD] = 11.23), with an overall range from 15 to 72 years. This means that the sample is rather young, which is in line with the recruitment channels and the survey being online-based (i.e., potentially less accessible for older people). Table 21 shows the complete distribution of age.

Table 21: Distribution of age.



Average household size is 2.45 (SD = 1.34), indicating that many participants live with family or roommates. Average product familiarity is 3.93 (SD = 0.80) on a 5-point Likert-type scale (1: Not at all familiar – 5: Extremely familiar), indicating rather high familiarity overall with products like the described Chocolate and Sweater. Table 22 summarizes the detailed descriptive results for age, household size and familiarity with the product.

Table 22: Descriptive results of numeric and interval scaled control variables.
Note: The number of total observations is 738, lower values for N in the table above indicate missing values in these control variables.

Variable	N	Mean	SD	Median	Minimum	Maximum	Skewness	Kurtosis
Age	736	32.02	11.23	28	15	72	1.33	1.20
Household	734	2.45	1.34	2	1	10	1.37	3.00
Familiar	735	3.93	0.80	4	1	5	-0.82	1.11

In terms of gender, slightly more women than man have participated (~54% versus ~45%), while seven participants identify as non-binary. Given the very small number of non-binary participants, these will not be analyzed as a separate group in the following analyses. Instead, men will be contrasted against female and non-binary participants combined. Almost 80% of participants have no children, which is in line with the rather young sample. About 70% of participants describe their living area as urban. 605 participants (~82%) have lived most of their lives in Germany, followed by 15 (~2%) participants in Türkiye and 13 (~2%) participants in Austria and USA each. No more than six participants lived longest in any other single country, respectively. Slightly more than 60% of participants have completed some kind of university degree. 231 participants (~35%) have a net household income of less than 25k € per year in the last year, while 68 participants (~10%) earn more than 99k € annually, and the remaining ~55% earn between 25k € and 99k € annually. Finally, ~70% of participants have chosen to answer the survey in German. Table 23 provides more details about the demographic composition of the sample.

Table 23: Descriptive results of nominal and ordinal scaled control variables.

Variable		N	%
Gender	Female	398	54%
	Male	333	45%
	Non-binary	7	1%
	Missing	0	0%
Children	< 5 years	48	7%
	5-9 years	26	4%
	10-14 years	17	2%
	15-19 years	14	2%
	> 19 years	51	7%
	No child	582	79%
	Missing	0	0%
Area	Rural	199	27%
	Urban	538	73%
	Missing	1	0%

Variable		N	%
Country	Germany	605	82%
	Other	132	18%
	Missing	1	0%
Education	Secondary school	175	24%
	Vocational / technical / trade training	107	14%
	Bachelor, Master, Diploma, Staatsexamen	427	58%
	Doctorate degree	28	4%
	Missing	1	0%
Income	< 25k €	231	31%
	25-49k €	180	24%
	50-74k €	118	16%
	75-99k €	64	9%
	> 99k €	68	9%
	Missing	77	10%
Survey language	English	220	30%
	German	518	70%
	Missing	0	0%

5.3.2. Data preparation and descriptive analyses

The final data sets per sub-sample were downloaded from Unipark (2024) on 04.12.2024, imported into R studio and consolidated into one comprehensive data set. Variables are given appropriate names to be easily identifiable and new variables are generated based on available data (e.g., a variable naming the product, or a variable identifying the age of the youngest child). The two reversed items are recoded to ensure scale consistency across all items. Reversed items had been included because they were part of scales adopted from previous research. The data is cleansed as described previously, including e.g., deleting observations with failed attention checks or with completion times of less than four minutes. Subsequently, the duration is analyzed (mean = 9.2 minutes; SD = 4.9 minutes; median = 8.0 minutes). Duration results for Chocolate and Sweater are highly similar, indicating that participants in both product scenarios paid the same amount of attention. Table 24 shows the details about the duration (for answering the questionnaire) for each product separately.

Table 24: Descriptive statistics: Duration (completion time) in minutes (per product).

Product	N	Mean	SD	Median	Minimum	Maximum
Chocolate	370	9.2	5.1	8.0	4.0	53.1
Sweater	368	9.2	4.7	8.0	4.0	41.3

After inspection of the control variables and duration, descriptive analyses of the key items commenced (via inspection of N, mean, median, SD, skewness, kurtosis). Several visualizations (e.g., boxplots, histograms, and Q-Q plots) are created and inspected to increase overall understanding of the data. During this analysis, the initially reversed item *Exper4* was found to show a bimodal distribution. Bimodal distributions indicate that the data consist of two underlying distinct groups (Siegel, 2012). Since this is a reverse item ("Its producer does not have much experience with eco-friendly products") in the questionnaire, it seems like one group of participants may have answered the reversed item and the other group may have overlooked the negation in the statement. In addition, it has been generally observed in marketing research, that reversed items can fit poorly or may distort the whole factor structure (Weijters & Baumgartner, 2012), especially if they are negation items (e.g., "…does not have much experience" in *Exper4*) as was the case in this study (Kam, 2023). To avoid further confusion or distortions, this item is excluded from the following analyses. *Effort2*, the other initially reversed item, shows a normal looking distribution for the products individually and combined, thus it is retained for further analyses.

Specifically for normality assessment, additional analyses are chosen: The skewness and kurtosis of all indicators are analyzed and a formal test is conducted. Skewness indicates how symmetric an indicator is distributed (negative skewness relates to a bigger number of high values in the answers), while kurtosis describes how the peak of the distribution is shaped (i.e., pointed or flat where negative kurtosis relates to a flat shape) (Eid et al., 2011). Values between -2 and +2 for both skewness and excess kurtosis are typically considered acceptable for normality (Hair et al., 2022). Most indicator values are within these thresholds. Second, Q-Q plots are used to visually inspect the distribution as they allow to compare the sample quantiles to quantiles as expected in case of normal distribution (Oppong & Agbedra, 2016). Third, as formal test, the Shapiro-Wilk test is conducted. The Shapiro-Wilk test was chosen due to its high performance for testing normality (Patrício et al., 2017; Razali & Wah, 2011; Shapiro et al., 1968; Yap & Sim, 2011). Inspection of skewness, kurtosis suggests normality for most items, Q-Q plots show some deviations, while the formal test indicates that none of the indicator items is normally distributed. However, many statistical analyses are quite robust against small deviations from normality, especially in case of large sample sizes (Eid et al., 2011). In addition, boxplots are created and inspected to detect outliers in the data. At this stage, no adjustments are made. Instead, after factor analysis and creation of mean scales, the analyses are repeated

for the then newly created measures. Similarly, standardization of variables is not applied at this stage but is revisited for the mean scales.

After the descriptive analyses and data cleansing where needed, the measurement scales are assessed via factor analysis (CFA) before hypotheses are tested via MLR and PLS SEM. The next sections explain why and how these methods are applied.

5.3.3. Factor analysis results

CFA is conducted using the lavaan package in R. First, two separate factor analyses are conducted for the two products. For the Chocolate, a hypothesized structure of 15 factors is tested (excluding Dur, Rep, and Rec because these dimensions are not applicable to Chocolate). For the Sweater, a hypothesized structure of 18 factors is assessed. Two random subsamples with each consisting of 80% of the observations are created from the full dataset of each product. This step is taken since some modifications are expected (e.g., derived from descriptive analyses) and it is recommended to cross-validate a modified CFA with a separate dataset to avoid overfitting to the specific data (Bühner, 2011; Weijters & Baumgartner, 2012). 80% of sample size was chosen because bigger samples commonly lead to more precise and reliable results in CFA (Gagne & Hancock, 2006), while the author also reasoned that this magnitude allows for enough variability to assess model fit after modification. Any modifications will be applied to both products to ensure consistency.

General measurement fit. The initial model for Chocolate already shows relatively good global fit to the data (scaled results using the Satorra-Bentler correction): χ^2 = 1,638.212, df = 1,169, $p \leq 0.001$, CFI = 0.937, TLI = 0.929, RMSEA = 0.037, SMSR = 0.059. The initial model for Sweater (based on 294 observations) shows a slightly less suitable but still reasonable global fit to the data (scaled results using the Satorra-Bentler correction): χ^2 = 2,334.443, df = 1,737, $p \leq 0.001$, CFI = 0.928, TLI = 0.919, RMSEA = 0.034, SMSR = 0.061. For both models, all standardized factor loadings are > 0.4. However, MIs show clear improvement potential: First, item *PQ1* has a very high cross-loading with PI. Given its phrasing ("This product is adapted to my needs."), it seems theoretically sound to assume that participants relate adaptation to their needs not so much to having high PQ in general but rather being relevant for themselves specifically and thus being willing / likely to purchase the product. Second, item *Soc3* ("It is more beneficial to society's welfare than other products.") shows high cross-loadings with multiple constructs. This indicates that it does not only measure Soc but rather can be interpreted in many ways in line with the broad term "society's welfare" (where society could be viewed as the consumer, as other people, as the natural environment, as animals, etc.).

Third, the item *Nat2* ("It has undergone minimal processing.") cross-loads especially with Res and Innov, indicating that participants might not relate the extent of being processed so much to Nat as to being overall efficient by not using resources during processing and as having innovative production processes.

Based on the statistical and theoretical considerations outlined, the three items *Nat2*, *Soc3*, and *PQ1* are excluded. Both models (Chocolate and Sweater) are iterated once using the same data and analyzed again. Global and local model fit is improved. Even though MIs remain, no further adjustments in terms of item exclusion are made to reduce the risk of overfitting, to retain more indicators rather than fewer and because theoretical justification for additional adjustments indicated by MIs seems less clear than for the previously adjusted. Analysis of construct reliability reveals that Rec does not meet any recommended thresholds for Cronbach alpha, McDonald omega and AVE. As a consequence, the construct Rec is dropped from further analysis, leaving it to be subject of future research. This decision has been taken because the multi-measure was self-developed and not previously validated and because it seems unreasonable to continue with a single item measure for this construct.

The iterated model (without *Nat2*, *Soc3*, *PQ1,* and Rec) using the second random data subsample, shows good fit for Chocolate: χ^2 = 1,273.660, df = 1,022, p ≤ 0.001, CFI = 0.964, TLI = 0.959, RMSEA = 0.029, SMSR = 0.046; and for Sweater: χ^2 = 1,592.126, df = 1,257, p ≤ 0.001, CFI = 0.955, TLI = 0.948, RMSEA = 0.030, SMSR = 0.052. Direct comparison of the initial and modified model shows a significant improvement both for Chocolate ($AIC_{initial}$ = 45,404, $AIC_{modifiied}$ = 42,484, $\Delta\chi^2$ = 345.13, Δdf = 147, p ≤ 0.001) and Sweater ($AIC_{initial}$ = 53,452, $AIC_{modifiied}$ = 44,901, $\Delta\chi^2$ = 744.20, Δdf = 480, p ≤ 0.001). An additional check with the complete observations of Chocolate and Sweater combined (N = 738), also shows a good fit: χ^2 = 1,388.780, df = 1,022, p ≤ 0.001, CFI = 0.978, TLI = 0.975, RMSEA = 0.022, SMSR = 0.037.

The resulting reliability indicators (i.e., Cronbach alpha, Omega and AVE) for the full sample are within recommended ranges with few exceptions: First, AVE_{Nat} = 0.46 (Cronbach $alpha_{Nat}$ = 0.75; McDonald $omega_{Nat}$ = 0.77); Second, AVE_{Res} = 0.48 and thus slightly below the recommended value of 0.5 (Cronbach $alpha_{Res}$ = 0.72; McDonald $omega_{Res}$ = 0.74), which is driven by the Sweater, while the values for Chocolate alone indicate good reliability (Chocolate only: Cronbach $alpha_{Res}$ = 0.76; McDonald $omega_{Res}$ = 0.77; AVE_{Res} = 0.54). Third and finally, AVE_{Effort} is slightly below 0.5 (Cronbach $alpha_{Effort}$ = 0.71; McDonald $omega_{Effort}$ = 0.72; AVE_{Effort} = 0.48). Given that the thresholds for Cronbach alpha and McDonald omega

are fulfilled for all constructs, no adjustments are made. Appendix C shows the final items used for further analyses based on the full model (all products and full sample), including measures of internal consistency reliability (Cronbach alpha and McDonald omega), and convergent validity (AVE and factor loading).

PEF dimensions. Beyond investigating and refining the consistency and convergent validity of all measures, CFA is used to analyze the hypothesized PEF dimensions for each product (except for Rec as described above). The latent reflective variable as measured by the dimensions is called PEF_dim in this thesis. Separate CFAs per product only focusing on PEF and the PEF dimensions (Nat, Soc, Ani, Cli, Res, Dur and Rep) are conducted. The selection of dimensions has been based on the results of pretest 1 (see section 5.2.1.1). To confirm the hypothesized structure and its relevance for PEF, a hierarchical model is built with the dimensions on the first level and PEF_dim as higher order construct on the second level. In addition, a regression between PEF_dim and PEF_uni is included to inspect the relation of PEF_dim with a different measure (indicating parallel test reliability). Individual results for both products are explained in the next paragraphs. Modelling is conducted based on the full dataset per product because no major modifications in terms of specifying indicators are expected at this stage.

The model for Chocolate contains the dimensions: Nat, Soc, Cli and Res, as these are the four dimensions considered important in pretest 1. The indicators as defined in the previous CFA analysis are included for analysis. The global model fit is good ($\chi^2 = 217.379$, df = 114, $p \leq 0.001$, CFI = 0.957, TLI = 0.948, RMSEA = 0.050, SMSR = 0.053). The correlation between PEF_dim and PEF_uni is 0.76, indicating good parallel test reliability. All indicators show significant standardized loadings > 0.4 to their respective factors. Thus, Nat, Soc, Cli and Res will be used as dimensions of Chocolate with all their indicators. Figure 14 visualizes the results of the CFA for Chocolate. Table 25 summarizes the standardized loadings on PEF_dim at the end of this section.

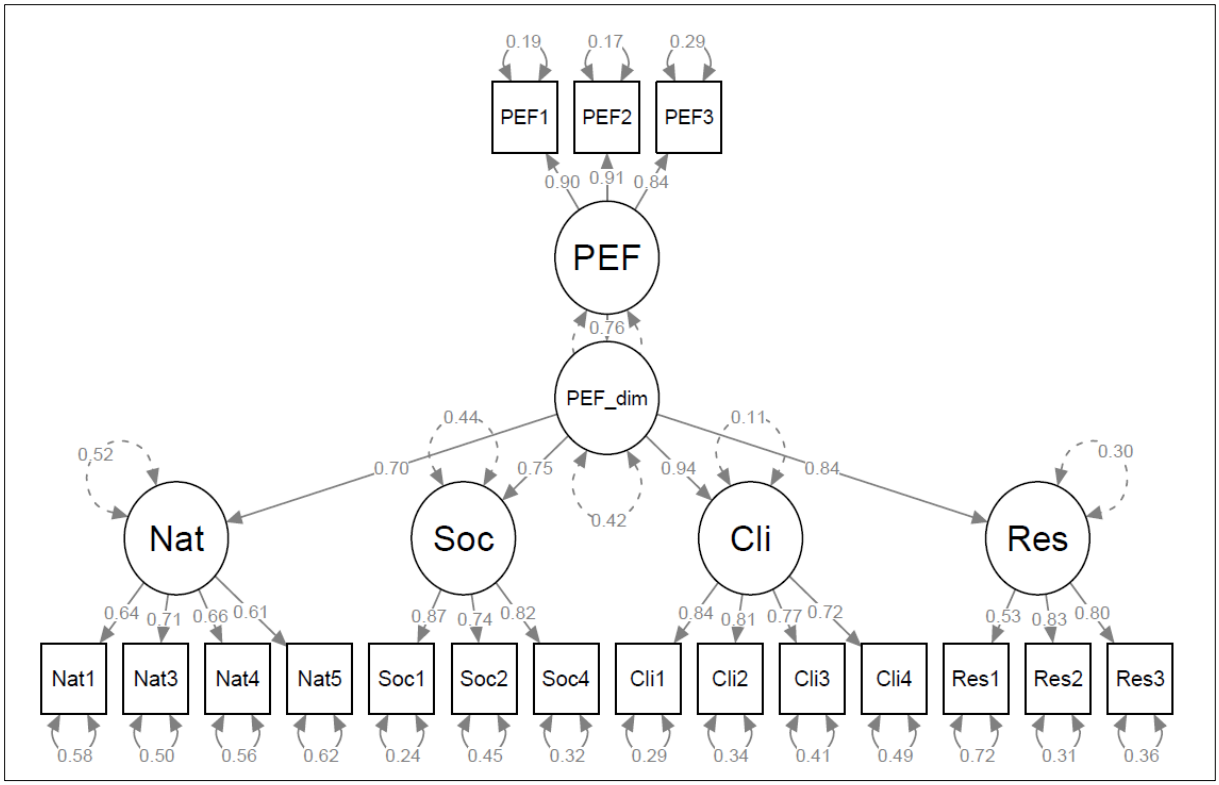


Figure 14: CFA results - PEF dimensions for Chocolate.

For the Sweater, all dimensions except Rec are selected as dimensions for further analysis (based on pretest 1 and CFA analysis as described above). The indicators for global fit indicate a reasonable model fit, however, Res is not a significant indicator of PEF_dim ($p > 0.05$). In addition, the standardized factor loading of Rep as indicator of PEF_dim is 0.28 and below the threshold of 0.4, even though it is significant ($p < 0.001$). Further inspection shows a high residual covariance between Rep and Dur indicating a close relationship between these two, which can theoretically be explained by both constructs being closely related to a longer product life. However, in this case, the contribution of Rep in excess of the other dimensions towards an overarching dimensional PEF is deemed negligible due to the very low factor loading. Thus, the model is specified without Res and Rep. The resulting global fit measures are: $\chi^2 = 270.484$, df = 183, $p \leq 0.001$, CFI = 0.968, TLI = 0.963, RMSEA = 0.036, SMSR = 0.065. All individual items and dimensions of PEF_dim show significant loadings to the respective factors (all standardized loadings above the threshold of 0.4). The correlation between PEF_dim and PEF_uni is 0.70, indicating parallel test reliability. Thus, Nat, Soc, Ani, Cli and Dur will be considered in the further analyses for the Sweater as PEF dimensions. Figure 15 visualizes the results of the CFA.

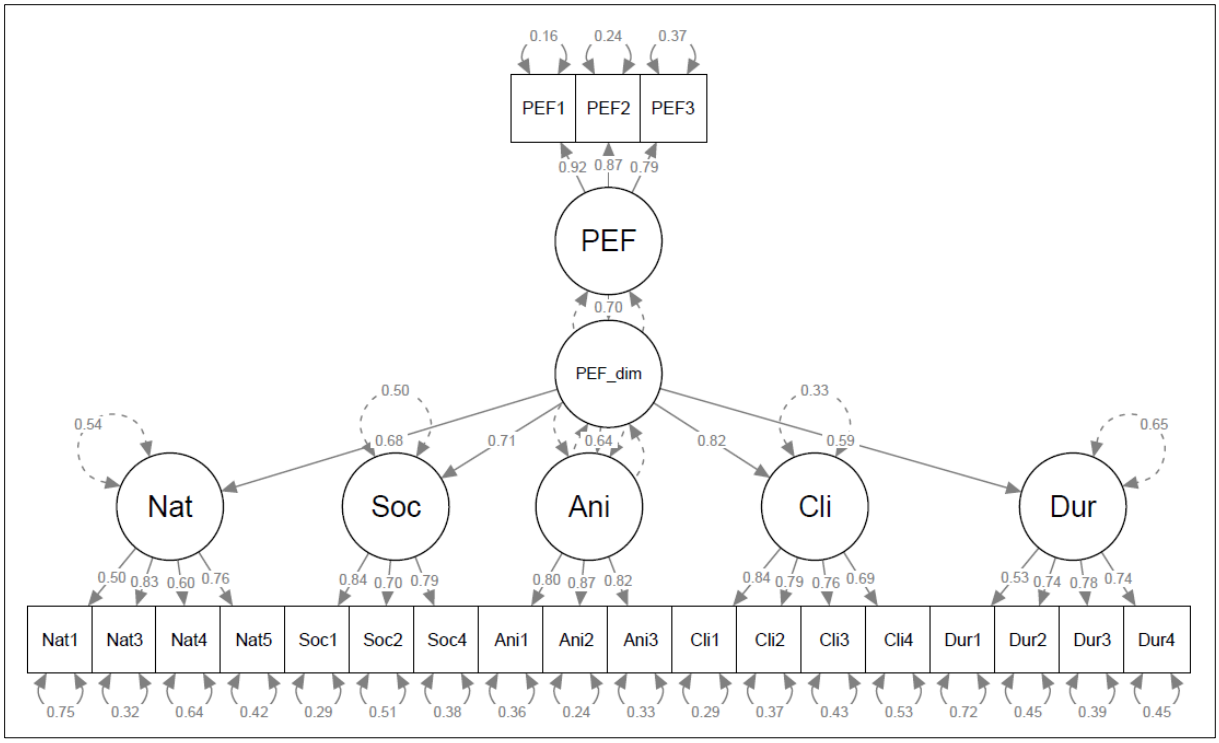


Figure 15: CFA results - PEF dimensions for Sweater.

When comparing Chocolate and Sweater, different dimensions are representing the dimensional PEF construct. Table 25 shows all eight initial dimensions as identified in the qualitative study and the standardized factor loadings for Chocolate and Sweater. It shows that the dimensions Nat, Soc and Cli are relevant for both Chocolate and Sweater, that Ani and Dur relevant only for the Sweater, while Res is relevant only for the Chocolate, and that Rep and Rec are not relevant for either product in this study. For Chocolate, all standardized factor loadings are ≥ 0.70, and for Sweater, all standardized factor loadings are ≥ 0.59.

Table 25: Standardized factor loadings of PEF dimensions.

	Standardized factor loadings	
Dimension	**Chocolate**	**Sweater**
Nat	0.70	0.68
Soc	0.75	0.71
Ani	-	0.64
Cli	0.94	0.82
Res	0.84	-
Dur	-	0.59
Rep	-	-
Rec	-	-

Summarizing the CFA results, the measurements in the questionnaire showed a good empirical fit to the theoretical setup after slight modifications and thus will be used as described above. For the analyses with PEF dimensions, the specific dimensions to be used per product have been confirmed. For the following regression analyses, mean scores are built, and specifically for the analyses with PEF dimensions, the mean score of PEF_dim is built by first creating mean scores of each applicable dimension and then averaging the respectively applicable dimensions again.

5.3.4. Regression results

MLR is used as the first step to 1) identify which control variables play a role and will be included in the final model, 2) identify if there are any InfObs that should be excluded and 3) identify if there are any hypothesized effects that should be excluded from the SEM analyses to reduce the model complexity. To conduct regression analyses, mean scores were built for all variables, because regression analysis does not incorporate latent variables and indicators (this will be addressed via SEM in the next section).

The regression analyses are conducted first considering PEF_uni (the unidimensional PEF measure) because this is an established way of representing PEF and will be used to test the hypothesized model. Afterwards, the analyses for PEF_dim (the multidimensional PEF measure) will be conducted and compared against the results based on the PEF_uni measure.

5.3.4.1. Selection of relevant control variables

For simplification purposes, control variables with more than three levels were translated into one dichotomous dummy variable per initial control variable: First, childrens' age was summarized into having no child versus at least one child of any age; Second, country of residence was summarized into "Germany" or "Other"; Third, Education was summarized into having any university degree (62% of observations) versus no university degree; and Fourth), Income was summarized into < 50k € annual household income (56% of observations) and ≥ 50k € annual household income. For income, around 50k € annual net income is in line with the average annual household net income in Germany (Kott, 2024). MLR with PQ as dependent variable and PEF_uni as predictor together with the collected control variables were analyzed. Values for PQ and PEF_uni were standardized, while the control variables were not transformed. Observations with missing values in the control variables were excluded since it could not be determined if the missing values occurred at random or not. This leads to a deviation from the number of observations, with only 328 observations for the Chocolate (instead of 370) and only 322 observations for the Sweater (instead of 368). The main variable

with missing values was income, where participants had been given the option to answer “prefer not to disclose”.

First, the assumptions for MLR are assessed for all models. For the combined model, errors are not normally distributed according to the Shapiro-Wilk test ($W = 0.993$, $p = 0.003$), while the Q-Q plots of residuals indicate no big deviations from normality. Autocorrelation could not be detected ($DW = 2.072$, p-value = 0.817), but heteroscedasticity ($BP = 29.540$, $p = 0.001$) so that robust standard errors are used. For the Chocolate model, errors are normally distributed ($W = 0.994$, $p = 0.286$), autocorrelation is not present ($DW = 2.023$, $p = 0.571$) and homoscedasticity is supported ($BP = 13.757$, $p = 0.184$). For the Sweater model, the Shapiro-Wilk test indicates that errors are not normally distributed ($W = 0.989$, $p = 0.012$), but visual inspection indicates no major deviations. There is no autocorrelation ($DW = 2.146$, $p = 0.901$), but heteroscedasticity ($BP = 20.925$, $p = 0.022$) which leads to using robust standard errors. Multicollinearity is not an issue in any of the models (all VIF values < 3). Table 26 shows the results of the regression models which are described in the following.

Table 26: MLR results investigating control variables.
Note: Number of observations lower than the full sample due to exclusion of observations with missing values in the control variables.

	Dependent variable: PQ		
	Chocolate	**Sweater**	**Combined**
	β (SE)	β (SE)	β (SE)
PEF_uni	0.468***	0.450***	0.463***
	(0.045)	(0.061)	(0.040)
Age	-0.007	-0.012	-0.009**
	(0.006)	(0.007)	(0.004)
Male	-0.142	-0.092	-0.121*
	(0.094)	(0.106)	(0.069)
No child	0.158	0.073	0.130
	(0.156)	(0.183)	(0.120)
Education: Uni	-0.170*	-0.029	-0.105
	(0.097)	(0.109)	(0.072)
Area: Urban	0.203*	0.032	0.116
	(0.104)	(0.116)	(0.076)
Income: < 50k €	0.038	-0.016	0.010
	(0.103)	(0.111)	(0.074)
Household	0.008	0.023	0.019
	(0.040)	(0.039)	(0.029)
Country_other	-0.229*	-0.249*	-0.231***
	(0.121)	(0.131)	(0.089)
Familiar	0.160***	0.073	0.116***
	(0.058)	(0.062)	(0.044)
Constant	-0.484	0.065	-0.234
	(0.452)	(0.477)	(0.320)
Observations	328	322	650
R^2	0.340	0.236	0.283
Adjusted R^2	0.319	0.211	0.271
Residual Std. Error	0.820 (df = 317)	0.901 (df = 311)	0.856 (df = 639)
F Statistic	16.318*** (df = 10; 317)	9.608*** (df = 10; 311)	25.170*** (df = 10; 639)

*p < 0.1; **p < 0.05; ***p < 0.01

In the model using both products, the adjusted R^2 is 0.27 ($F(10; 639) = 25.17$, $p < 0.001$), higher values of PEF_uni increased PQ ($\beta_{PEF\_uni} = 0.46$, $p < 0.001$), and being younger ($\beta_{Age} = -0.01$, $p < 0.040$), more familiar with the product ($\beta_{Familiar} = 0.12$, $p = 0.004$), or living in Germany for most of one's life ($\beta_{Country\_other} = -0.23$, $p = 0.089$), leads to higher PQ as compared to being less familiar, older or having lived longest in a country that is not Germany. No other control variable is significant. For the Sweater, no control variable shows a statistically significant effect. PEF_uni significantly increases PQ ($\beta_{PEF\_uni} = 0.45$, $p < 0.001$) in the model that explains 21% of variance ($F(10,311) = 9.61$, $p < 0.001$). For the Chocolate, again PEF_uni ($\beta_{PEF\_uni} = 0.47$, $p < 0.001$) and familiarity ($\beta_{Familiar} = 0.16$, $p = 0.007$) increase PQ. Within this model ($R^2 = 0.32$, $F(10,317) = 16.32$, $p < 0.001$), familiarity is also significant ($\beta_{Familiar} = 0.16$, $p = 0.007$). Only for Chocolate, an additional quadratic relationship between PEF_uni and PQ is found, which however does not change the significance of the control variables and only has a minor impact on the coefficient of familiarity ($\beta_{Familiar} = 0.17$, $p = 0.04$) and a minor impact on the overall model ($R^2 = 0.33$, $F(11,316) = 15.41$, $p < 0.001$). Upon further investigation, the non-linear effect within the Chocolate model is eliminated if outliers in PEF_uni with values of PEF_uni < 3 are excluded from the analyses, meaning that for most cases only the linear relationship plays a role. At this time, no outliers are removed because InfObs for the regression models are analyzed in the next step in more detail. Since the impact on control variables is negligible, the choice of control variables to be included in further analyses is not impacted by the non-linear effect.

For the following analyses, only familiarity, age, gender (as dummy coded variable comparing being male against female and non-binary combined) and country of longest residence (as dummy coded variable comparing Germany versus other) are included because these variables showed a significant impact in the combined product model and no other control variables had a significant impact in any of the three models.

Observations with missing values in the control variables used for further analyses (N = 6) are excluded from the subsequent analyses (MLR and PLS SEM) results because it is unclear if they occur at random. As next step, the impact of potentially influential variables is assessed with the goal of deciding if any and which observations should be excluded further).

5.3.4.2. Selection of final sample based on influential observations

As described in section 5.2.3.2, InfObs should be investigated and decisions taken about their inclusion or exclusion. The analyses for InfObs were conducted for both products separately to avoid any distortions by mixing products together. Identification of InfObs was conducted with functionalities from the packages car (Fox & Weisberg, 2019) and stats (R Core Team, 2024) in R. These packages have built-in functionalities that compute and visualize studentized residuals, leverage / hat values, (S)DFBETA, (S)DFFIT, Cook's distance and COVRATIO values. The model used to identify InfObs uses PQ as dependent variable and all hypothesized constructs as predictors and is formalized as follows: PQ ~ PEF + Wglow + Innov + Stren + Indic + Effort + Exper (+ Age + Male + Familiar + Country). Note that the analysis was conducted twice: With and without the control variables shown in brackets.

All models (Chocolate and Sweater with and without control variables) have normally distributed residuals, show no autocorrelation, and no concerning multicollinearity (all VIF values < 3), but they show heteroscedasticity. Heteroscedasticity leads to inaccurate standard errors and can thus impact some indicators for InfObs like Cook's distance. To assess the potential impact of inaccurate standard errors, the standard errors outputted by the regression formula are compared with robust standard error (as derived from the heteroscedasticity-consistent covariance matrix using White's estimator (H. White, 1980)). The differences in standard errors of all variables are very small, for example for the Sweater model with control variables the average difference between "normal" and robust standard errors across all variables is < 0.0007. This finding in combination with the fact that identification of InfObs is never an exact science, leads to the decision to evaluate measures like Cook's distance regardless of heteroscedasticity.

Including the control variables into the analyses of InfObs reduces the number of InfObs from 42 (19 for Chocolate and 23 for Sweater) to 24 (13 for Chocolate and 11 for Sweater), emphasizing the relevance of the control variables. For comparison purposes, a third way of identifying InfObs is deployed: All hypothesized direct effects with PEF and PQ as dependent variables are assessed in simple linear regressions individually (without control variables) and InfObs identified – all observations that are considered influential in more than half of the separate linear regressions for either PEF or PQ are then defined as influential for the whole analysis.

After the identification of potentially InfObs, the same regression model as indicated above is estimated again with two differences: First, the hypothesized interaction effects are added into the model. Second, different data are used to compare results: (1) all observations, (2) without InfObs identified controlling for age, gender, familiarity and country, (3) without InfObs identified without control variables, and (4) without InfObs identified from individual regressions (e.g., PQ ~ Exper, PEF_uni ~ Exper). Table 27 and Table 28 show the results for Chocolate and Sweater, respectively. All results described in the following are based on the standardized regression coefficients for the hypothesized effects of the latent constructs (for best comparability between constructs) and unstandardized coefficients for the control variables (for better interpretability, e.g., age in years is easier to understand than a standardized version).

Before interpreting the results, the assumptions for regressions are inspected. For the Chocolate, visual inspection of residuals via respective Q-Q plots indicates normal distributed errors for all models, even though the Shapiro-Wilk test is significant for models (2) and (3) ($W_{(1)} = 0.998$, $p_{(1)} = 0.880$; $W_{(2)} = 0.991$, $p_{(2)} = 0.039$; $W_{(3)} = 0.990$, $p_{(3)} = 0.028$; $W_{(4)} = 0.995$, $p_{(4)} = 0.279$). Because the deviations from normality are small, this is not considered an issue for the current data analysis. All models show heteroscedasticity via the Breusch-Pagan test ($BP_{(1)} = 50.41$, $p_{(1)} < 0.001$; $BP_{(2)} = 28.13$, $p_{(2)} = 0.009$; $BP_{(3)} = 27.83$, $p_{(3)} = 0.010$; $BP_{(4)} = 31.34$, $p_{(4)} = 0.003$) and their standard errors are adjusted for interpretation of significance levels. No model has autocorrelation of errors indicated by the Durbin Watson test ($DW_{(1)} = 2.026$, $p_{(1)} = 0.586$; $DW_{(2)} = 1.965$, $p_{(2)} = 0.361$; $DW_{(3)} = 2.006$, $p_{(3)} = 0.511$; $DW_{(4)} = 2.030$, $p_{(4)} = 0.599$) and multicollinearity is not concerning in any model (all VIF values < 3).

Table 27: MLR results comparing models with differing observations for Chocolate. Notes: (1) all observations, (2) without InfObs identified controlling for age, gender, familiarity and country, (3) without InfObs identified without control variables, and (4) without InfObs identified from individual regressions.

	Dependent variable: PQ			
	(1)	**(2)**	**(3)**	**(4)**
	β (SE)	β (SE)	β (SE)	β (SE)
PEF_uni	0.326***	0.372***	0.348***	0.377***
	(0.058)	(0.056)	(0.056)	(0.063)
Wglow	0.081	0.093	0.112*	0.082
	(0.066)	(0.066)	(0.066)	(0.065)
Innov	-0.008	-0.028	-0.030	0.019
	(0.057)	(0.050)	(0.050)	(0.054)
Stren	-0.084*	-0.079*	-0.081*	-0.101**
	(0.044)	(0.044)	(0.044)	(0.046)
Indic	0.039	0.047	0.031	0.032
	(0.046)	(0.047)	(0.047)	(0.047)
Effort	0.093**	0.078*	0.091**	0.107**
	(0.046)	(0.045)	(0.046)	(0.047)
Exper	0.216***	0.223***	0.251***	0.185***
	(0.054)	(0.057)	(0.059)	(0.051)
Age	-0.010***	-0.008**	-0.007**	-0.009***
	(0.003)	(0.003)	(0.003)	(0.003)
Male	-0.137	-0.146*	-0.164**	-0.148*
	(0.084)	(0.081)	(0.081)	(0.082)
Familiar	0.089*	0.054	0.060	0.082
	(0.057)	(0.054)	(0.054)	(0.052)
Country_other	-0.226**	-0.247**	-0.232**	-0.199*
	(0.110)	(0.109)	(0.104)	(0.107)
PEF_uni:Stren	-0.017	0.031	0.054	0.032
	(0.052)	(0.049)	(0.048)	(0.058)
PEF_uni:Indic	0.037	0.017	0.021	0.060
	(0.058)	(0.056)	(0.055)	(0.063)
Constant	0.071	0.158	0.102	0.041
	(0.274)	(0.260)	(0.257)	(0.261)
Observations	367	354	348	351
R^2	0.399	0.423	0.429	0.398
Adjusted R^2	0.377	0.401	0.407	0.375
Residual Std.	0.709	0.751	0.742	0.759
Error	(df = 353)	(df = 340)	(df = 334)	(df = 337)
F Statistic	18.026*** (df = 13; 353)	19.163*** (df = 13; 340)	19.311*** (df = 13; 334)	17.134*** (df = 13; 337)

*p < 0.1; **p < 0.05; ***p < 0.01

Analyzing the results for the Chocolate, the model based on the full dataset explains ~38% of the variance in PQ ($F(13,353) = 18.03$, $p < 0.01$). The explained variance varies between 38% (model 1 and model 4) and 41% (model 3). The residual standard errors do not vary much between models, within a range of 0.709 (model 1) and 0.759 (model 4). Thus, even though excluding InfObs seems to slightly modify results, the differences between models are very small. Formal assessment via the Akaike criterion shows that model 3 has the highest fit: $AIC_{(1)} = 805$, $AIC_{(2)} = 818$, $AIC_{(3)} = 796$, $AIC_{(4)} = 819$. This is not surprising because more InfObs have been excluded from that model than from the others. Comparing this "best-fitting" model 3 with the original model including all observations, there are no differences regarding the significance levels of any construct – the only difference is found for the control variable gender, which has become significant in model 3 ($\beta_{Male} = -0.16$, $p = 0.043$) as compared to model 1 ($\beta_{Male} = -0.14$, $p = 0.120$). The impact of PEF_uni on PQ has become slightly stronger as indicated by a higher regression coefficient (model 3: $\beta_{PEF\_uni} = 0.35$, $p < 0.001$ versus model 1: $\beta_{PEF\_uni} = 0.33$, $p < 0.001$). The potential non-linear effect of PEF_uni for Chocolate was analyzed and is not significant in any of the models, thus it will not be further considered in the analyses (as it was not hypothesized).

For the Sweater, the assumptions for regressions are also inspected before interpreting the results. Visual inspection of residuals via Q-Q plots indicate normally distributed errors for all models, even though the Shapiro-Wilk test is significant for models 2 and 3 ($W_{(1)} = 0.993$, $p_{(1)} = 0.114$; $W_{(2)} = 0.987$, $p_{(2)} = 0.003$; $W_{(3)} = 0.989$, $p_{(3)} = 0.012$; $W_{(4)} = 0.992$, $p_{(4)} = 0.068$). This is not considered an issue for the current data analysis because the deviations from normality are rather small. All models are considered heteroscedastic via the Breusch-Pagan test ($BP_{(1)} = 42.44$, $p_{(1)} < 0.001$; $BP_{(2)} = 42.65$, $p_{(2)} < 0.001$; $BP_{(3)} = 39.40$, $p_{(3)} < 0.001$; $BP_{(4)} = 43.35$, $p_{(4)} < 0.001$), thus their standard errors are adjusted for interpretation of significance levels. No model has autocorrelation of errors as indicated by the Durbin Watson test ($DW_{(1)} = 2.102$, $p_{(1)} = 0.832$; $DW_{(2)} = 2.155$, $p_{(2)} = 0.925$; $DW_{(3)} = 2.155$, $p_{(3)} = 0.923$; $DW_{(4)} = 2.157$, $p_{(4)} = 0.930$) and multicollinearity is not found in any model (all VIF values < 3).

Table 28: MLR results comparing models with differing observations for Sweater. Notes: (1) all observations, (2) without InfObs identified controlling for age, gender, familiarity and country, (3) without InfObs identified without control variables, and (4) without InfObs identified from individual regressions.

	Dependent variable: PQ			
	(1)	**(2)**	**(3)**	**(4)**
	β (SE)	β (SE)	β (SE)	β (SE)
PEF_uni	0.288***	0.319***	0.297***	0.308***
	(0.061)	(0.055)	(0.056)	(0.059)
Wglow	0.149**	0.164***	0.239***	0.124**
	(0.068)	(0.069)	(0.071)	(0.067)
Innov	0.054	0.056	0.027	0.060
	(0.057)	(0.056)	(0.057)	(0.057)
Stren	-0.084*	-0.054	-0.089*	-0.046
	(0.048)	(0.046)	(0.048)	(0.047)
Indic	0.203***	0.211***	0.195***	0.196***
	(0.054)	(0.052)	(0.057)	(0.055)
Effort	0.007	0.005	-0.025	0.016
	(0.047)	(0.047)	(0.048)	(0.046)
Exper	-0.024	-0.025	-0.052	0.007
	(0.057)	(0.056)	(0.058)	(0.053)
Age	-0.011**	-0.008**	-0.005	-0.006
	(0.004)	(0.004)	(0.004)	(0.004)
Male	-0.084	-0.081	-0.064	-0.073
	(0.091)	(0.087)	(0.088)	(0.090)
Familiar	0.030	0.037	0.056	0.034
	(0.054)	(0.055)	(0.055)	(0.056)
Country_other	-0.085	-0.089	-0.117	-0.133
	(0.119)	(0.120)	(0.119)	(0.115)
PEF_uni:Stren	0.023	0.045	0.050	0.008
	(0.051)	(0.044)	(0.045)	(0.046)
PEF_uni:Indic	-0.029	-0.038	-0.043	-0.005
	(0.053)	(0.049)	(0.054)	(0.055)
Constant	0.286	0.197	0.036	0.147
	(0.284)	(0.278)	(0.282)	(0.280)
Observations	365	354	342	354
R^2	0.315	0.359	0.367	0.309
Adjusted R^2	0.290	0.335	0.342	0.283
Residual Std. Error	0.844 (df = 351)	0.797 (df = 340)	0.781 (df = 328)	0.806 (df = 340)
F Statistic	12.412*** (df = 13; 351)	14.667*** (df = 13; 340)	14.639*** (df = 13; 328)	11.694*** (df = 13; 340)

*p < 0.1; **p < 0.05; ***p < 0.01

Analyzing the separate models for the Sweater, the differences between models seem to be similarly small like for the Chocolate, with explained variances ranging from $R^2_{(4)} = 0.28$ to $R^2_{(2)} = 0.34$. Comparing models via AIC values shows the best fit for model 3: $AIC_{model\ 1} = 928$, $AIC_{model\ 2} = 860$, $AIC_{model\ 3} = 817$, $AIC_{model\ 4} = 868$. Significance levels of the hypothesized constructs do not show any relevant variation between models. Like for Chocolate, the effects of the independent variables vary only minimally between the initial model 1 and the "best fitting" model 3, e.g., $\beta_{PEF\_uni} = 0.29$ in model 1, $p < 0.001$ versus $\beta_{PEF\_uni} = 0.30$, $p < 0.001$ in model 3. Age is the only variable where significance changes between the models ($\beta_{Age} = -0.011$ in model 1, $p = 0.014$ versus $\beta_{Age} = -0.005$, $p = 0.163$ in model 3).

To ensure applicability, the same analyses as described above were conducted for the models considering PEF_dim instead of PEF_uni. Results confirmed the suitability of the respective "model 1" for Chocolate and Sweater (see Appendix D and E for details). Given the results presented in this section, the author of this thesis decided not to exclude any observations from the subsequent analyses because no thorough theoretical reasoning could be derived and the statistical differences were rather small. Thus, the results from the respective "model 1" for Chocolate and Sweater are interpreted.

5.3.4.3. Selection of hypotheses

To gain first insights into the relevance of specific hypotheses and identify with effects could be excluded to increase parsimony, MLR regressions with PQ, PEF and PI respectively as dependent variables were conducted.

MLR with PQ as dependent variable. For Chocolate, the overall model is significant ($R^2 = 0.38$, $F(13,353) = 18.03$, $p < 0.001$). As expected, PEF_uni shows the strongest effect ($\beta_{PEF\_uni} = 0.33$, $p < 0.001$) together with Exper ($\beta_{Exper} = 0.22$, $p < 0.001$). Effort ($\beta_{Effort} = 0.09$, $p = 0.044$) also positively impacts PQ. Higher Stren ($\beta_{Stren} = -0.08$, $p = 0.057$) fail to produce a significant impact on PQ, as does Wglow ($\beta_{Wglow} = 0.08$, $p = 0.223$). Innov ($\beta_{Innov} = -.01$, $p = 0.886$), and Indic ($\beta_{Indic} = 0.04$, $p = 0.392$), are not relevant for PQ, and neither are the interaction effects between Stren ($\beta_{PEF\_uni:Stren} = -0.02$, $p = 0.744$) and Indic ($\beta_{PEF\_uni:Indic} = 0.04$, $p = 0.526$) with PEF_uni. Age is significant as socio-demographic variable ($\beta_{Age} = -0.01$, $p = 0.003$), indicating that being older is related to lower PQ, and having lived outside of Germany for most of one's life is also related to lower PQ ($\beta_{Country\_other} = -0.23$, $p = 0.043$), while the other control variables have no significant impact on PQ ($\beta_{Male} = -0.14$, $p = 0.106$; $\beta_{Familiar} = 0.09$, $p = 0.120$).

For the Sweater, the picture in terms of which independent variables are significant looks quite different than for Chocolate. The overall model is significant ($R^2 = 0.29$, $F(13,351) = 12.41$, $p < 0.001$). PEF_uni has by far the strongest effect ($\beta_{PEF\_uni} = 0.29$, $p < 0.001$), followed by Indic ($\beta_{Indic} = 0.20$, $p < 0.001$). Wglow ($\beta_{Wglow} = 0.15$, $p = 0.028$) has the only remaining significant effect of the hypothesized constructs. For the Sweater, Stren ($\beta_{Stren} = -0.08$, $p = 0.083$), Effort ($\beta_{Effort} = 0.01$, $p = 0.873$) and Exper ($\beta_{Exper} = -0.03$, $p = 0.669$) are not relevant for PQ, and neither are the interaction effects between Stren ($\beta_{PEF\_uni:Stren} = 0.02$, $p = 0.651$) and Indic ($\beta_{PEF\_uni:Indic} = -0.03$, $p = 0.581$) with PEF_uni. Age is the only significant socio-demographic variable ($\beta_{Age} = -0.01$, $p = 0.014$), yet with a very small impact. However, it indicates that being older is related to lower PQ across products. No other control variable is significant at the 5% level ($\beta_{Country\_other} = -0.08$, $p = 0.476$; $\beta_{Male} = -0.08$, $p = 0.354$; $\beta_{Familiar} = 0.03$, $p = 0.578$).

Overall, the purpose of the regression described in this section was to determine the suitable dataset in terms of observations to be excluded or included and get a first impression of the relevance of the hypothesized effects towards PQ. It was decided not to exclude any observations due to lacking theoretical justification combined with small statistical effects. The respective models with PQ as dependent variable were then interpreted in terms of effects of the independent variables. Chocolate and Sweater showed a vastly different pattern in terms of which constructs have a significant impact. What was common to both was the clearly positive impact of PEF_uni on PQ, the lack of evidence for the interaction effects (between Stren and Indic with PEF_uni) and the (small) negative impact of age towards PQ evaluation.

MLR with PEF as dependent variable. Following the analyses regarding the hypothesized effects on PQ, the hypothesized effects on PEF_uni are analyzed next to identify the relevant control variables impacting PEF_uni. In terms of assumptions for MLR, visual inspection of residuals via respective Q-Q plots indicates mostly normal distributed errors despite deviations at the lower end of the curve for Chocolate. The Shapiro-Wilk test is significant for both products ($W_{Chocolate} = 0.983$, $p_{Chocolate} < 0.001$; $W_{Sweater} = 0.988$, $p_{Sweater} = 0.004$). The Breusch-Pagan test indicates heteroscedasticity for Chocolate but not Sweater ($BP_{Chocolate} = 37.92$, $p_{Chocolate} < 0.001$; $BP_{Sweater} = 13.30$, $p_{Sweater} = 0.065$) so that standard errors are adjusted for interpretation of significance levels of Chocolate. Given the adjustment of standard errors for Chocolate and that these regressions are only an intermediate step towards analyzing the full model, the slight deviation from normality of errors is not considered critical at this stage. There is no autocorrelation of errors according to the Durbin Watson test

($DW_{Chocolate} = 1.982$, $p_{Chocolate} = 0.423$; $DW_{Sweater} = 1.947$, $p_{Sweater} = 0.298$) and multicollinearity is not concerning in either model (all VIF values < 3). Table 29 shows the summarized results of the regressions.

Table 29: MLR results with (unidimensional) PEF as dependent variable.

	Dependent variable: PEF_uni	
	Chocolate	**Sweater**
	β (SE)	β (SE)
Innov	0.172***	0.180***
	(0.054)	(0.051)
Effort	0.089**	0.080
	(0.045)	(0.049)
Exper	0.384***	0.328***
	(0.064)	(0.054)
Age	0.002	-0.001
	(0.004)	(0.004)
Male	-0.021	0.149
	(0.091)	(0.095)
Familiar	0.113*	0.079
	(0.066)	(0.062)
Country_other	0.132	-0.112
	(0.108)	(0.124)
Constant	-0.518*	-0.322
	(0.312)	(0.307)
Observations	367	365
R^2	0.281	0.223
Adjusted R^2	0.267	0.208
Residual Std. Error	0.859 (df = 359)	0.887 (df = 357)
F Statistic	20.028*** (df = 7; 359)	14.626*** (df = 7; 357)

*p < 0.1; **p < 0.05; ***p < 0.01

Both models are significant ($F_{Chocolate}(7,359) = 28.88$, $p < 0.001$; $F_{Sweater}(7,357) = 27.90$, $p < 0.001$).The explained variance is with 27% higher for the Chocolate compared to 21% explained in the Sweater scenario. Regarding the impact of the independent variables, results show that Innov is a significant predictor of PEF_uni in both product scenarios (Chocolate: $\beta_{Innov} = 0.17$, $p = 0.002$; Sweater: $\beta_{Innov} = 0.18$, $p < 0.001$), as well as Exper (Chocolate: $\beta_{Exper} = 0.38$, $p < 0.001$; Sweater: $\beta_{Exper} = 0.33$, $p < 0.001$). Effort has a positive impact for Chocolate ($\beta_{Effort} = 0.09$, $p = 0.046$), while it not significant for the Sweater ($\beta_{Effort} = 0.08$, $p = 0.104$). In

terms of control variables, being more familiar with the product is only directionally significant for the Chocolate ($\beta_{Familiar} = 0.11$, $p = 0.089$), but not for the Sweater ($\beta_{Familiar} = 0.08$, $p = 0.200$). The other included control variables have no statistically relevant impact.

Results for the MLR with PEF_dim are overall similar. Q-Q plots indicate mostly normal distributed errors, while the Shapiro-Wilk tests are significant ($W_{Chocolate} = 0.988$, $p_{Chocolate} = 0.004$; $W_{Sweater} = 0.992$, $p_{Sweater} = 0.038$). The Breusch-Pagan test indicates heteroscedasticity ($BP_{Chocolate} = 23.98$, $p_{Chocolate} = 0.001$; $BP_{Sweater} = 29.95$, $p_{Sweater} < 0.001$) so that standard errors are adjusted. The Durbin Watson test indicates no autocorrelation of errors ($DW_{Chocolate} = 1.900$, $p_{Chocolate} = 0.164$; $DW_{Sweater} = 1.865$, $p_{Sweater} = 0.094$) and multicollinearity is not concerning (all VIF values < 3). The MLR models are significant for both Chocolate and Sweater ($F_{Chocolate}(7,359) = 28.88$, $F_{Sweater}(7,357) = 27.90$). The explained variance is with 34% higher for Sweater than for Chocolate, where it is 28%. Regarding the impact of the independent variables, results show that Innov is a significant predictor of PEF_dim (Chocolate: $\beta_{Innov} = 0.16$, $p = 0.006$; Sweater: $\beta_{Innov} = 0.23$, $p < 0.001$), as well as Exper (Chocolate: $\beta_{Exper} = 0.40$, $p < 0.001$; Sweater: $\beta_{Exper} = 0.42$, $p < 0.001$). Effort here is not significant for Chocolate ($\beta_{Effort} = 0.08$, $p = 0.129$), while it has a positive impact for the Sweater ($\beta_{Effort} = 0.11$, $p = 0.012$). In terms of control variables, being more familiar with the product relates to higher PEF_dim for the Chocolate ($\beta_{Familiar} = 0.16$, $p = 0.014$), but not for the Sweater ($\beta_{Familiar} = 0.09$, $p = 0.073$) alone. No other control variable included in the model is statistically relevant.

The regression results with PEF (measured uni- and multidimensionally) as dependent variable indicate overall the relevance of the hypothesized effects and suggest familiarity with the product as additional influence, while no other control variables seem to impact PEF. Thus, for the SEM, only familiarity will be considered as control variable towards PEF.

MLR with PI as dependent variable. Beyond the investigated effects towards PQ and PEF, the hypothesized model considers significant impact of PQ and Wglow towards PI. This is analyzed in a model with PI as dependent variable, PQ and Wglow as predictors and the same control variables as previously identified. As in the other regressions, standardized values for the theoretical constructs are used. Inspection of the assumptions indicates a non-normal distribution of errors via the Shapiro-Wilk test ($W_{Chocolate} = 0.956$, $p_{Chocolate} < 0.001$; $W_{Sweater} = 0.976$, $p_{Sweater} < 0.001$) and visual inspection shows a strong deviation at the lower end of the curve in the Q-Q plots. Heteroscedasticity ($BP_{Chocolate} = 5.36$, $p_{Chocolate} = 0.498$; $BP_{Sweater} = 7.99$, $p_{Sweater} = 0.239$), autocorrelation ($DW_{Chocolate} = 1.967$, $p_{Chocolate} = 0.366$; $DW_{Sweater} = 1.962$,

$p_{Sweater}$ = 0.352) and multicollinearity (all VIF values < 3) raise no concerns. Thus, due to lacking normality, bootstrapping with 5,000 resamples is conducted to estimate the coefficients (and standard errors) and their 95% CI. Results are shown in Table 30 and described in the following.

Table 30: MLR results with PI as dependent variable (bootstrapped).
Note: Estimates written in bold are statistically significant (p < 0.05).

	Dependent variable: PI			
	Chocolate		**Sweater**	
	Estimate β (SE)	95% CI	Estimate β (SE)	95% CI
PQ	**0.283** (0.055)	**[0.171, 0.388]**	**0.283** (0.056)	**[0.173, 0.340]**
Wglow	**0.229** (0.061)	**[0.115, 0.351]**	**0.232** (0.062)	**[0.111, 0.351]**
Age	0.003 (0.004)	[-0.005, 0.010]	0.003 (0.004)	[-0.005, 0.010]
Male	-0.106 (0.096)	[-0.311, 0.073]	-0.116 (0.098)	[-0.317, 0.069]
Familiar	**0.191** (0.062)	**[0.073, 0.316]**	**0.187** (0.061)	**[0.074, 0.312]**
Country_other	0.040 (0.129)	[-0.244, 0.272]	0.039 (0.130)	[-0.249, 0.266]
Constant	-0.811 (0.316)	[-1.435, -0.198]	-0.795 (0.313)	[-1.419, -0.202]

Bootstrapping results demonstrate a consistent view of the relationships with PI across products. For Chocolate, higher PQ (β_{PQ} = 0.28, 95% CI [0.17, 0.39]), and Wglow (β_{Wglow} = 0.23, 95% CI [0.12, 0.35]) both lead to higher PI. In terms of control variables, increased familiarity with the product also impacts PI positively ($\beta_{Familiar}$ = 0.19, 95% CI [0.07, 0.32]), while age (β_{Age} = 0.00, 95% CI [-0.01, 0.01]), gender (β_{Male} = -0.11, 95% CI [-0.31, 0.07]) and country of residence ($\beta_{Country\_other}$ = 0.04, 95% CI [-0.24, 0.27]) have no effects. The Chocolate model explains 24% of the variance in PI as calculated via the original model without bootstrapping. For the Sweater, bootstrapped results are nearly identical, even in magnitude of individual effects, while the overall model explains 36% of variance (β_{PQ} = 0.28, 95% CI [0.17, 0.34]; β_{Wglow} = 0.23, 95% CI [0.11, 0.35]; $\beta_{Familiar}$ = 0.19, 95% CI [0.07, 0.31]; β_{Age} = 0.00, 95% CI [-0.01, 0.01]; β_{Male} = -0.12, 95% CI [-0.32, 0.07]; $\beta_{Country\_other}$ = 0.04, 95% CI [-0.25, 0.27]).

In this section, the relationships with PI have been assessed and both hypothesized variables (PQ and Wglow) show a significant positive impact. In addition, higher familiarity with the product increases PI. Thus, in the following SEM analyses, familiarity will also be considered towards PI.

In addition to these reported results, individual regressions assessing the main effects were conducted separately for each hypothesis as introduced in section 4.3.4. The results show that each hypothesized predictor shows a significant main effect with PQ regardless of the product used, providing further support for the relevance of the identified constructs from qualitative research. However, as observed in the MLR regression with PQ as dependent variable, the interaction effects fail to have a significant impact. Thus, these will not be further considered in the PLS SEM analyses. Based on the regression analyses as reported in this section, only selected control variables are considered for the SEM analyses: Age, gender, familiarity, and country of residence towards PQ; and familiarity towards PEF and PI. The next section explains the results of the SEM analyses in detail.

5.4. Structural equation modelling results

This section describes the results of the hypotheses investigations via SEM analyses. PLS SEM was conducted using the seminr package (Ray et al., 2024) in R and 10,000 bootstrap iterations per model. Figure 16 shows the final model under investigation, based on the initial hypotheses and the refinement via regression analyses as described in the last section. Two main modifications were made to the previous theoretical model: First, the initially hypothesized interaction effects (Indic and Stren with PEF, respectively) were taken out due to lack of empirical support. Second, specific control variables were selected based on the regression analyses (as described in the last sections) and are shown in the model. Age, gender, familiarity with the product und country of longest residence (Germany versus others) were selected to control PQ. In addition, familiarity with the product was selected to control PEF and PI.

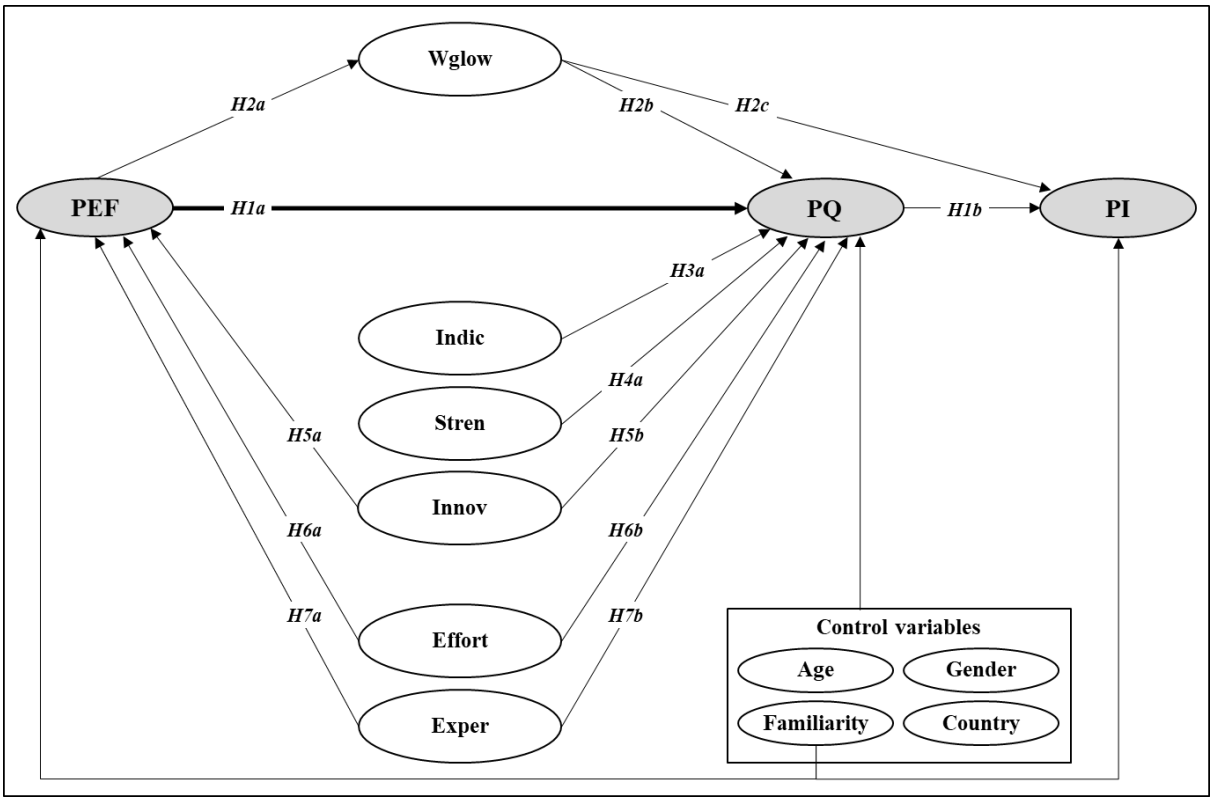


Figure 16: Final model under investigation via SEM.

This model was assessed four times in total with different data: First considering the unidimensional measure for PEF and analyzing the Chocolate and Sweater separately. Afterwards, the same analyses were repeated considering PEF as higher-order dimensional construct – again for Chocolate and for Sweater separately. The following sections first present the results of considering PEF_uni, then of PEF_dim, and finally compare all models (across PEF measures and products).

5.4.1. Results from the unidimensional PEF measure

This section addresses the key research question of the quantitative part of this thesis: Which of the proposed constructs influence the relations between perceived eco-friendliness (PEF) and perceived quality (PQ)? It describes first the results from the measurement models for Chocolate and Sweater and then for the structural model for Chocolate, followed by the structural model for Sweater.

5.4.1.1. Results from measurement models

Before describing and interpreting the results from the different structural models, the measurement models were assessed via factor loadings, Cronbach alpha, rhoC, AVE, HTMT and FLC as described in section 5.2.2.1. Confirming the findings from the previous CFA, all

indicators in all SEM analyses showed significant factor loadings > 0.4. Assessment of HTMT and FLC showed that all models fulfilled the criteria for discriminant validity: All HTMT values were < 0.80 and the construct correlations of a given construct were smaller than the square root of the AVE of that construct in all models. In addition, multicollinearity was not concerning in any model as indicated by VIF values below the threshold of 10. In fact, all VIF values were < 3. Thus, the criteria for reliability and validity are confirmed overall and the structural models can be interpreted.

Table 31 shows all values for Cronbach alpha, rhoC and AVE across the models analyzing Chocolate and Sweater separately. As expected based on the previous CFA results, reliability and convergent validity were within the recommended thresholds with the exception of one construct: The AVE for Effort was < 0.5. For Effort and Sweater as product, rhoC was also slightly below the recommended threshold of 0.7 while Cronbach alpha was exactly at the threshold level. Thus, the measurement of Effort is to some extent lacking reliability and validity. However, it will be retained in the analysis, yet results need to be interpreted carefully and should be reconfirmed in the future, for example with an adjusted measure of Effort.

Table 31: Reliability and convergent validity of measures used across PLS SEM (unidimensional PEF).

	Chocolate			Sweater		
Construct	alpha	rhoC	AVE	alpha	rhoC	AVE
PEF_uni	0.914	0.914	0.780	0.893	0.894	0.738
Indic	0.758	0.787	0.571	0.790	0.805	0.588
Effort	0.709	0.713	0.455	0.700	0.672	0.418
Exper	0.900	0.900	0.751	0.887	0.888	0.726
Stren	0.873	0.869	0.576	0.885	0.879	0.603
Innov	0.843	0.845	0.648	0.831	0.827	0.617
Wglow	0.863	0.863	0.678	0.906	0.906	0.763
PQ	0.903	0.905	0.761	0.896	0.895	0.741
PI	0.947	0.947	0.857	0.898	0.899	0.747

5.4.1.2. Results from the structural model for Chocolate

After having assessed the suitability of the measures used in the PLS SEM analyses, their results in terms of explained variances of endogenous constructs, path coefficients and path significances are described. Table 32 shows the results of the structural model for Chocolate with the unidimensional PEF measure, which are explained in the following.

Table 32: PLS SEM results for Chocolate (unidimensional PEF).
Notes: Paths written in bold are statistically significant as indicated by the 95% confidence interval (p < 0.05); n.s. = not significant.

Hypo-thesis	Path			Estimate (β)	Boot-strap SD	T Statistic	2.5% CI	97.5% CI	f^2	Significance
1a	**PEF_uni**	**→**	**PQ**	**0.356**	**0.083**	**4.286**	**0.189**	**0.515**	**0.113**	**significant**
1b	**PQ**	**→**	**PI**	**0.320**	**0.067**	**4.798**	**0.185**	**0.446**	**0.106**	**significant**
2a	**PEF_uni**	**→**	**Wglow**	**0.706**	**0.036**	**19.805**	**0.633**	**0.772**	**0.993**	**significant**
2b	Wglow	→	PQ	0.087	0.120	0.724	-0.142	0.328	0.004	n.s.
2c	**Wglow**	**→**	**PI**	**0.295**	**0.071**	**4.159**	**0.162**	**0.440**	**0.092**	**significant**
3a	Indic	→	PQ	0.041	0.060	0.686	-0.078	0.159	0.002	n.s.
4a	**Stren**	**→**	**PQ**	-0.100	0.053	-1.874	-0.207	0.001	0.016	n.s.
5a	**Innov**	**→**	**PEF_uni**	**0.164**	**0.067**	**2.445**	**0.029**	**0.293**	**0.033**	**significant**
5b	Innov	→	PQ	-0.024	0.077	-0.316	-0.174	0.128	0.001	n.s.
6a	Effort	→	PEF_uni	0.088	0.058	1.519	-0.026	0.200	0.011	n.s.
6b	**Effort**	**→**	**PQ**	**0.118**	**0.060**	**1.960**	**0.002**	**0.240**	**0.023**	**significant**
7a	**Exper**	**→**	**PEF_uni**	**0.448**	**0.066**	**6.758**	**0.323**	**0.579**	**0.235**	**significant**
7b	**Exper**	**→**	**PQ**	**0.204**	**0.069**	**2.940**	**0.071**	**0.343**	**0.043**	**significant**
n/a	**Age**	**→**	**PQ**	**-0.129**	**0.042**	**-3.100**	**-0.213**	**-0.048**	**0.031**	**significant**
n/a	Male	→	PQ	-0.071	0.045	-1.585	-0.155	0.019	0.009	n.s.
n/a	**Germany**	**→**	**PQ**	**0.090**	**0.046**	**1.954**	**0.002**	**0.184**	**0.015**	**significant**
n/a	**Familiar**	**→**	**PEF_uni**	**0.109**	**0.054**	**2.028**	**0.005**	**0.216**	**0.018**	**significant**
n/a	Familiar	→	PQ	0.076	0.049	1.562	-0.017	0.172	0.010	n.s.
n/a	Familiar	→	PI	0.143	0.049	2.914	0.044	0.237	0.030	n.s.

The model with Chocolate as product and using PEF measured unidimensionally explains between 30% and 50% of variances in the endogenous constructs: $R^2_{PEF\_uni} = 0.35$, $R^2_{PQ} = 0.47$, $R^2_{Wglow} = 0.50$, $R^2_{PI} = 0.35$. Results show high consistency with the results achieved from the MLR models in the previous section. Specifically, it is found that PEF_uni is positively and significantly related to PQ ($\beta_{PEF\_uni\text{-}PQ} = 0.36$, 95% CI [0.19, 0.52]) and to Wglow ($\beta_{PEF\_uni\text{-}Wglow} = 0.71$, 95% CI [0.63, 0.77]). Contrary to the initial expectation, Wglow is not found to be a mediator between PEF_uni and PQ, as shown by the non-significant path from Wglow to PQ ($\beta_{Wglow\text{-}PQ} = 0.09$, 95% CI [-0.14, 0.33]). Wglow does however show a positive impact on PI ($\beta_{Wglow\text{-}PI} = 0.30$, 95% CI [0.16, 0.44]). PQ is found to be positively impacted by Effort ($\beta_{Effort\text{-}PQ} = 0.12$, 95% CI [0.00, 0.24]) and by Exper ($\beta_{Exper\text{-}PQ} = 0.20$, 95% CI [0.07, 0.34]) in addition to PEF_uni. Additionally, age shows a significant impact towards PQ ($\beta_{Age\text{-}PQ} = -0.13$, 95% CI [-0.21, -0.05]), indicating that being older leads to lower PQ, and also country of residence, indicating that people having lived most of their lives in Germany show higher PQ ($\beta_{Germany\text{-}PQ}$

= -0.09, 95% CI [0.00, 0.18]). Both PQ and Wglow impact PI as expected ($\beta_{PQ\text{-}PI}$ = 0.32, 95% CI [0.19, 0.45]; $\beta_{Wglow\text{-}PI}$ = 0.30, 95% CI [0.16, 0.44]).

In addition to assessing the path estimates and significances, the predictive value of the model (for predicting PQ and PI) was assessed via ten-fold cross-validation, repeated ten times. The analysis focused on PQ and PI as these two constructs are the focal dependent variables. For both PQ and PI, the RMSE is lower for the PLS SEM as compared with the respective MLR model (see Appendix F for details) – indicating a high predictive power for these constructs.

5.4.1.3. Results from the structural model for Sweater

After having analyzed the results for Chocolate, the same was done for the Sweater. Table 33 shows the results which are described in the following.

Table 33: PLS SEM results for Sweater (unidimensional PEF).
Notes: Paths written in bold are statistically significant as indicated by the 95% confidence interval (p < 0.05); n.s. = not significant.

Hypo-thesis	Path			Estimate (β)	Boot-strap SD	T Statistic	2.5% CI	97.5% CI	f^2	Significance
1a	**PEF_uni**	**→**	**PQ**	**0.367**	**0.089**	**4.106**	**0.188**	**0.508**	**0.126**	**significant**
1b	**PQ**	**→**	**PI**	**0.400**	**0.053**	**7.597**	**0.296**	**0.501**	**0.223**	**significant**
2a	**PEF_uni**	**→**	**Wglow**	**0.629**	**0.039**	**16.209**	**0.546**	**0.698**	**0.654**	**significant**
2b	Wglow	→	PQ	0.053	0.106	0.498	-0.110	0.270	0.002	n.s.
2c	**Wglow**	**→**	**PI**	**0.304**	**0.058**	**5.235**	**0.185**	**0.414**	**0.129**	**significant**
3a	**Indic**	**→**	**PQ**	**0.269**	**0.080**	**3.375**	**0.125**	**0.409**	**0.077**	**significant**
4a	Stren	→	PQ	-0.114	0.065	-1.757	-0.214	0.018	0.017	n.s.
5a	**Innov**	**→**	**PEF_uni**	**0.258**	**0.153**	**1.685**	**0.126**	**0.382**	**0.072**	**significant**
5b	Innov	→	PQ	0.086	0.120	0.715	-0.077	0.247	0.006	n.s.
6a	**Effort**	**→**	**PEF_uni**	**0.132**	**0.219**	**0.604**	**0.031**	**0.273**	**0.022**	**significant**
6b	Effort	→	PQ	0.022	0.122	0.183	-0.124	0.139	0.001	n.s.
7a	**Exper**	**→**	**PEF_uni**	**0.338**	**0.221**	**1.527**	**0.202**	**0.470**	**0.109**	**significant**
7b	Exper	→	PQ	-0.050	0.087	-0.569	-0.191	0.101	0.002	n.s.
n/a	**Age**	**→**	**PQ**	**-0.097**	**0.048**	**-2.008**	**-0.185**	**-0.004**	**0.016**	**significant**
n/a	Male	→	PQ	-0.053	0.053	-1.015	-0.164	0.036	0.004	n.s.
n/a	Germany	→	PQ	0.038	0.052	0.730	-0.064	0.136	0.002	n.s.
n/a	Familiar	→	PEF_uni	0.041	0.082	0.508	-0.054	0.121	0.002	n.s.
n/a	Familiar	→	PQ	0.031	0.044	0.697	-0.047	0.124	0.001	n.s.
n/a	**Familiar**	**→**	**PI**	**0.208**	**0.046**	**4.544**	**0.117**	**0.297**	**0.076**	**significant**

The Sweater model using the unidimensional PEF measure explains between 31% and 44% of variances in the endogenous constructs: $R^2_{PEF\_uni}$ = 0.31, R^2_{PQ} = 0.39, R^2_{Wglow} = 0.39,

R^2_{PI} = 0.44. Results show high consistency with the results achieved from the MLR models in the previous section. Specifically, it is found that PEF_uni is positively and significantly related to PQ ($\beta_{PEF\_uni\text{-}PQ}$ = 0.37, 95% CI [0.19, 0.51]) and to Wglow ($\beta_{PEF\_uni\text{-}Wglow}$ = 0.63, 95% CI [0.55, 0.70]). Contrary to the initial expectation, Wglow is not significantly impacting PQ ($\beta_{Wglow\text{-}PQ}$ = 0.05, 95% CI [-0.11, 0.27]), while it does significantly impact PI ($\beta_{Wglow\text{-}PI}$ = 0.30, 95% CI [0.19, 0.41]). PQ is found to be positively impacted by Indic ($\beta_{Indic\text{-}PQ}$ = 0.27, 95% CI [0.13, 0.41]), but not by any other of the hypothesized constructs: $\beta_{Stren\text{-}PQ}$ = -0.11, 95% CI [-0.21, 0.02], $\beta_{Innov\text{-}PQ}$ = 0.09, 95% CI [-0.08, 0.25], $\beta_{Effort\text{-}PQ}$ = 0.02, 95% CI [-0.12, 0.14], $\beta_{Exper\text{-}PQ}$ = -0.05, 95% CI [-0.19, 0.10]. As expected, both PQ and Wglow impact PI ($\beta_{PQ\text{-}PI}$ = 0.40, 95% CI [0.30, 0.50]; $\beta_{Wglow\text{-}PI}$ = 0.30, 95% CI [0.19, 0.41]).

Out of the control variables, age shows a significant impact towards PQ ($\beta_{Age\text{-}PQ}$ = -0.10, 95% CI [-0.19, -0.00]), indicating that being older leads to lower PQ. Gender ($\beta_{Male\text{-}PQ}$ = -0.05, 95% CI [-0.16, 0.04]), country of residence ($\beta_{Germany\text{-}PQ}$ = 0.04, 95% CI [-0.06, 0.14]) and familiarity with the product ($\beta_{Familiar\text{-}PQ}$ = 0.03, 95% CI [-0.05, 0.12]) seem not to be relevant for PQ.

The predictive value of the model (for predicting PQ and PI) was assessed for the Sweater via ten-fold cross-validation, repeated ten times. Like for Chocolate, the RMSE for all indicators of PI is lower for PLS SEM as compared to the respective MLR model. For PQ, one indicator (*PQ2*) shows a higher RMSE value for the PLS SEM out of sample metrics than for the MLR out of sample metrics (see Appendix G for details). This indicates a medium to high predictive power for these constructs.

5.4.2. Dimensionality of PEF

After having addressed the primary research question by presenting the results of the SEM analyses and individual paths when considering PEF_uni, this section focuses on the proposed multidimensional definition of PEF (PEF_dim). This section individually explains the results for Chocolate and Sweater before the next section then compares the models. Like in the previous SEM analyses, all results are based on bootstrapping with 10,000 iterations. PEF_dim is operationalized as higher-order construct PEF_dim, which for Chocolate is represented by the dimensions Nat, Soc, Cli and Res, and for Sweater is represented by the dimensions Nat, Soc, Ani, Cli, and Dur. To incorporate the higher-order construct into the model, the disjoint two stage approach as described by Hair et al. (2024) is applied (explained in section 5.2.3.3). In short, there is a first stage where the latent variable scores of the

dimensions are calculated via SEM and then added as new variables to the dataset so that in the second stage, those variable scores are treated as observed variables indicating PEF_dim.

5.4.2.1. Results from measurement models

Before describing and interpreting the results from the different structural models, the measurement models were assessed via Cronbach alpha, rhoC, AVE, FLC and HTMT. Suitability for analyses and interpretation as indicated by the FLC and HTMT was given for both models. Table 34 shows all values for Cronbach alpha, rhoC and AVE across the models analyzing Chocolate and Sweater separately.

Table 34: Reliability and convergent validity of measures used across PLS SEM (multidimensional PEF).

	Chocolate			Sweater		
Construct	alpha	rhoC	AVE	alpha	rhoC	AVE
PEF_dim	0.827	0.830	0.553	0.783	0.782	0.422
Indic	0.758	0.787	0.571	0.790	0.805	0.588
Effort	0.709	0.713	0.457	0.700	0.679	0.423
Exper	0.900	0.900	0.750	0.887	0.887	0.724
Stren	0.873	0.869	0.576	0.885	0.879	0.604
Innov	0.843	0.844	0.646	0.831	0.829	0.618
Wglow	0.863	0.863	0.678	0.906	0.906	0.763
PQ	0.903	0.904	0.760	0.896	0.895	0.741
PI	0.947	0.947	0.857	0.898	0.899	0.747

As expected based on the previous CFA results, reliability and convergent validity were within the recommended thresholds with two exceptions: First, PEF_dim for the Sweater showed an AVE value below 0.5, indicating that the construct explains less than 50% variance of its indicators (i.e., PEF dimensions). While this could be taken as a hint towards a formative definition of PEF or a misspecification of the lower order constructs underlying PEF_dim, the measurement logic will not be adjusted within this thesis but should be revisited from a theoretical perspective in the future. Despite failing the recommended threshold for AVE (and for rhoC), the Sweater model based on PEF_dim will be investigated, compared and contrasted with the same model using the unidimensional measure (which fulfils all reliability and validity criteria). Second, the AVE for Effort was < 0.5 in all scenarios. This has already been found in the CFA results and in the measurement model analysis of the models with the unidimensional PEF measure. Like before, even though the measurement of Effort is lacking reliability and

validity to some extent, it will be retained in the analyses. Results will need to be interpreted carefully and should be reconfirmed with an adjusted measure of Effort at a later time.

5.4.2.2. Results from the structural model for Chocolate

For Chocolate, PEF_dim was represented by Nat, Soc, Cli and Res. Results from PLS SEM (with 10,000 bootstrap iterations) show explained variances within the range of 0.35 and 0.64 ($R^2_{PEF\_dim} = 0.47$, $R^2_{PQ} = 0.64$, $R^2_{Wglow} = 0.45$, $R^2_{PI} = 0.35$). The full results are shown in Table 35 and explained in the following.

Table 35: PLS SEM results for Chocolate (multidimensional PEF).
Notes: Paths written in bold are statistically significant as indicated by the 95% confidence interval ($p < 0.05$); n.s. = not significant.

Hypo-thesis	Path			Estimate (β)	Boot-strap SD	T Statistic	2.5% CI	97.5% CI	f^2	Significance
1a	**PEF_dim**	**→**	**PQ**	**0.352**	**0.139**	**2.531**	**0.106**	**0.653**	**0.070**	**significant**
1b	**PQ**	**→**	**PI**	**0.321**	**0.068**	**4.746**	**0.185**	**0.449**	**0.107**	**significant**
2a	**PEF_dim**	**→**	**Wglow**	**0.798**	**0.034**	**23.307**	**0.727**	**0.861**	**1.748**	**significant**
2b	Wglow	→	PQ	0.064	0.160	0.399	-0.285	0.338	0.002	n.s.
2c	**Wglow**	**→**	**PI**	**0.294**	**0.071**	**4.139**	**0.158**	**0.434**	**0.092**	**significant**
3a	Indic	→	PQ	0.015	0.063	0.241	-0.106	0.138	0.000	n.s.
4a	**Stren**	**→**	**PQ**	**-0.122**	**0.056**	**-2.182**	**-0.237**	**-0.015**	**0.023**	**significant**
5a	**Innov**	**→**	**PEF_dim**	**0.186**	**0.072**	**2.576**	**0.034**	**0.319**	**0.051**	**significant**
5b	Innov	→	PQ	-0.019	0.086	-0.216	-0.181	0.156	0.000	n.s.
6a	Effort	→	PEF_dim	0.069	0.060	1.150	-0.044	0.191	0.008	n.s.
6b	**Effort**	**→**	**PQ**	**0.129**	**0.061**	**2.126**	**0.011**	**0.249**	**0.026**	**significant**
7a	**Exper**	**→**	**PEF_dim**	**0.526**	**0.072**	**7.264**	**0.392**	**0.675**	**0.393**	**significant**
7b	**Exper**	**→**	**PQ**	**0.191**	**0.090**	**2.122**	**0.005**	**0.358**	**0.034**	**significant**
n/a	**Age**	**→**	**PQ**	**-0.141**	**0.043**	**-3.257**	**-0.221**	**-0.054**	**0.035**	**significant**
n/a	Male	→	PQ	-0.080	0.046	-1.724	-0.171	0.012	0.011	n.s.
n/a	Germany	→	PQ	0.050	0.047	1.054	-0.042	0.145	0.004	n.s.
n/a	**Familiar**	**→**	**PEF_dim**	**0.129**	**0.055**	**2.368**	**0.020**	**0.234**	**0.031**	**significant**
n/a	Familiar	→	PQ	0.068	0.050	1.350	-0.028	0.169	0.008	n.s.
n/a	**Familiar**	**→**	**PI**	**0.143**	**0.050**	**2.867**	**0.044**	**0.240**	**0.030**	**significant**

Like for the unidimensional measure, results were again consistent with the results from regression analyses: Higher levels of PEF_dim increase PQ ($\beta_{PEF\_dim\text{-}PQ} = 0.35$, 95% CI [0.11, 0.65]) and Wglow ($\beta_{PEF\_dim\text{-}Wglow} = 0.80$, 95% CI [0.73, 0.86]). Again, both PQ ($\beta_{PQ\text{-}PI} = 0.32$, 95% CI [0.19, 0.45]) and Wglow ($\beta_{Wglow\text{-}PI} = 0.29$, 95% CI [0.16, 0.43]) are positively related with PI. Effort ($\beta_{Effort\text{-}PQ} = 0.13$, 95% CI [0.01, 0.25]), Exper ($\beta_{Exper\text{-}PQ} = 0.19$, 95% CI [0.00,

0.36]), and Stren ($\beta_{Stren\text{-}PQ}$ = -0.12, 95% CI [-0.24, -0.02]), show a positive impact on PQ, while Indic ($\beta_{Indic\text{-}PQ}$ = 0.02, 95% CI [-0.11, 0.14]), Innov ($\beta_{Innov\text{-}PQ}$ = -0.02, 95% CI [-0.18, 0.16]), and Wglow ($\beta_{Wglow\text{-}PQ}$ = 0.06, 95% CI [-0.29, 0.34]), have no statistically significant influence on PQ. Regarding PEF_dim as endogenous construct, higher Innov ($\beta_{Innov\text{-}PEF\_dim}$ = 0.19, 95% CI [0.03, 0.32]), and Exper ($\beta_{Exper\text{-}PEF\_dim}$ = 0.53, 95% CI [0.39, 0.68]), raise PEF_dim, while Effort ($\beta_{Effort\text{-}PEF\_dim}$ = 0.07, 95% CI [-0.04, 0.19]) does not. Regarding the control variables, age influences PQ ($\beta_{Age\text{-}PQ}$ = -0.14, 95% CI [-0.22, -0.05]) and familiarity influences PEF_dim ($\beta_{Familiar\text{-}PEF\_dim}$ = 0.13, 95% CI [0.02, 0.23]) and PI ($\beta_{Familiar\text{-}PI}$ = 0.14, 95% CI [0.04, 0.24]).

Figure 17 visualizes the path coefficients within the model, showing results for the unidimensional and the multidimensional PEF measures (before and after the slash, respectively). When considering PEF_dim, almost all path coefficients are at least as high or higher compared to using PEF_uni. For Stren, the impact on PQ is only significant in the model based on PEF_dim.

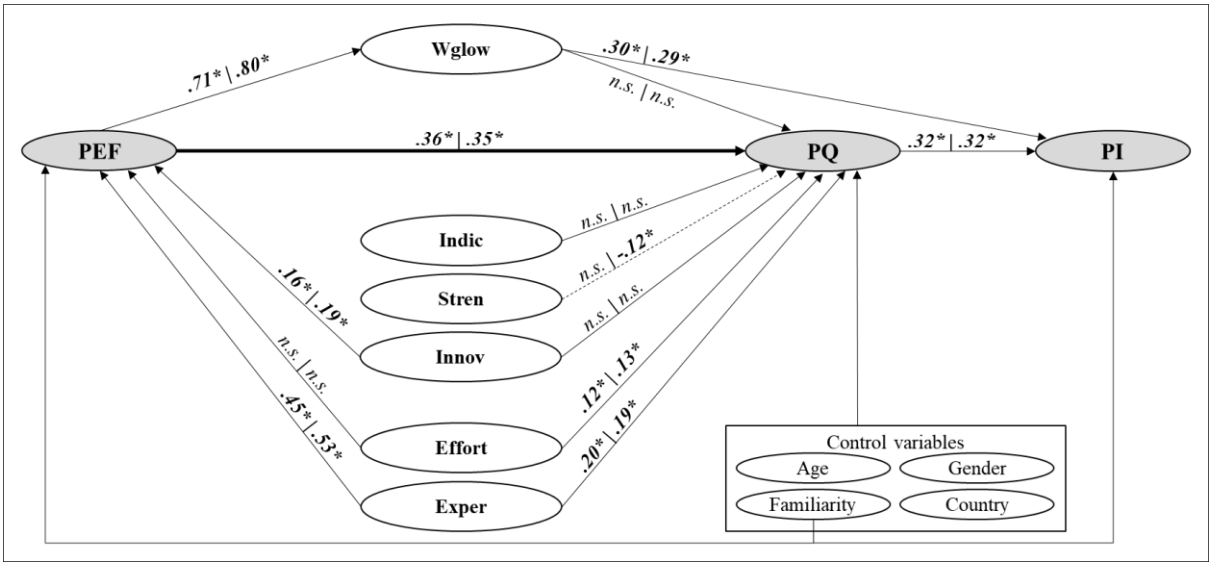


Figure 17: Visualization of SEM results for Chocolate.
*Notes: Values before the slash show beta coefficients from the model with PEF_uni, and values after the slash beta coefficients from the model with PEF_dim. Results for control variables are not visualized here. *p < 0.05*

5.4.2.3. Results from the structural model for Sweater

For Sweater, Nat, Soc, Ani, Cli and Dur represented PEF_dim. Results from PLS SEM (with 10,000 bootstrap iterations) show explained variances within the range of 0.43 and 0.55 ($R^2_{PEF\_dim} = 0.55$, $R^2_{PQ} = 0.43$, $R^2_{Wglow} = 0.52$, $R^2_{PI} = 0.44$). Table 36 shows the details before results are described in the following.

Table 36: PLS SEM results for Sweater (multidimensional PEF).
Notes: Paths written in bold are statistically significant as indicated by the 95% confidence interval ($p < 0.05$); n.s. = not significant.

Hypo-thesis	Path			Estimate (β)	Boot-strap SD	T Statistic	2.5% CI	97.5% CI	f^2	Significance
1a	**PEF_dim**	**→**	**PQ**	**0.620**	**0.143**	**4.338**	**0.396**	**0.965**	**0.209**	**significant**
1b	**PQ**	**→**	**PI**	**0.399**	**0.053**	**7.526**	**0.296**	**0.501**	**0.223**	**significant**
2a	**PEF_dim**	**→**	**Wglow**	**0.719**	**0.041**	**17.706**	**0.640**	**0.797**	**1.071**	**significant**
2b	Wglow	→	PQ	0.014	0.112	0.124	-0.214	0.213	-0.000	n.s.
2c	**Wglow**	**→**	**PI**	**0.305**	**0.059**	**5.192**	**0.184**	**0.414**	**0.130**	**significant**
3a	Indic	→	PQ	0.108	0.085	1.271	-0.065	0.265	0.012	n.s.
4a	Stren	→	PQ	-0.061	0.062	-0.986	-0.178	0.062	0.005	n.s.
5a	**Innov**	**→**	**PEF_dim**	**0.277**	**0.074**	**3.766**	**0.158**	**0.397**	**0.126**	**significant**
5b	Innov	→	PQ	0.065	0.096	0.678	-0.119	0.236	0.004	n.s.
6a	**Effort**	**→**	**PEF_dim**	0.135	0.101	1.336	**0.014**	**0.261**	**0.035**	**significant**
6b	Effort	→	PQ	0.019	0.080	0.236	-0.115	0.149	0.001	n.s.
7a	**Exper**	**→**	**PEF_dim**	**0.517**	**0.094**	**5.511**	**0.379**	**0.643**	**0.398**	**significant**
7b	**Exper**	**→**	**PQ**	**-0.195**	**0.087**	**-2.236**	**-0.394**	**-0.053**	**0.032**	**significant**
n/a	**Age**	**→**	**PQ**	**-0.121**	**0.050**	**-2.400**	**-0.215**	**-0.027**	**0.026**	**significant**
n/a	Male	→	PQ	-0.064	0.050	-1.298	-0.163	0.033	0.007	n.s.
n/a	Germany	→	PQ	0.070	0.057	1.225	-0.038	0.185	0.008	n.s.
n/a	Familiar	→	PEF_dim	0.085	0.066	1.294	-0.003	0.173	0.016	n.s.
n/a	Familiar	→	PQ	0.027	0.048	0.568	-0.063	0.114	0.001	n.s.
n/a	**Familiar**	**→**	**PI**	**0.207**	**0.046**	**4.545**	**0.116**	**0.294**	**0.076**	**significant**

PEF_dim positively impacts PQ ($\beta_{PEF\_dim\text{-}PQ} = 0.62$, 95% CI [0.40, 0.96]) and Wglow ($\beta_{PEF\_dim\text{-}Wglow} = 0.72$, 95% CI [0.64, 0.80]). Both PQ ($\beta_{PQ\text{-}PI} = 0.40$, 95% CI [0.30, 0.50]) and Wglow ($\beta_{Wglow\text{-}PI} = 0.31$, 95% CI [0.18, 0.41]) are positively related to PI. Exper shows a negative influence on PQ ($\beta_{Exper\text{-}PQ} = -0.20$, 95% CI [-0.39, -0.05]), which is contradictory with the initial hypothesis, which assumed a positive influence. However, Exper does positively influence PEF_dim ($\beta_{Exper\text{-}PEF\_dim} = 0.52$, 95% CI [0.38, 0.64]). Wglow ($\beta_{Wglow\text{-}PQ} = 0.01$ 95% CI [-0.21, 0.21]), Indic ($\beta_{Indic\text{-}PQ} = 0.11$, 95% CI [-0.07, 0.27]), Stren ($\beta_{Stren\text{-}PQ} = -0.06$, 95% CI [-0.18, 0.06]), Innov ($\beta_{Innov\text{-}PQ} = 0.07$, 95% CI [-0.12, 0.24]), and Effort ($\beta_{Effort\text{-}PQ} = 0.02$, 95%

CI [-0.12, 0.15]), have no statistically significant influence on PQ. Regarding PEF_dim as endogenous construct, higher Innov ($\beta_{Innov\text{-}PEF\_dim}$ = 0.28, 95% CI [0.16, 0.40]) increases PEF_dim. In terms of control variables, age influences PQ ($\beta_{Age\text{-}PQ}$ = -0.12, 95% CI [-0.22, -0.03]) and familiarity influences PI ($\beta_{Familiar\text{-}PI}$ = 0.21, 95% CI [0.12, 0.29]).

Figure 18 visualizes the path coefficients within the model, showing results for PEF_uni and PEF_dim (before and after the slash, respectively). When using PEF_dim, almost all path coefficients are at least as high or higher in absolute value compared to PEF_uni, with the exception of Indic (not significant when using PEF_dim).

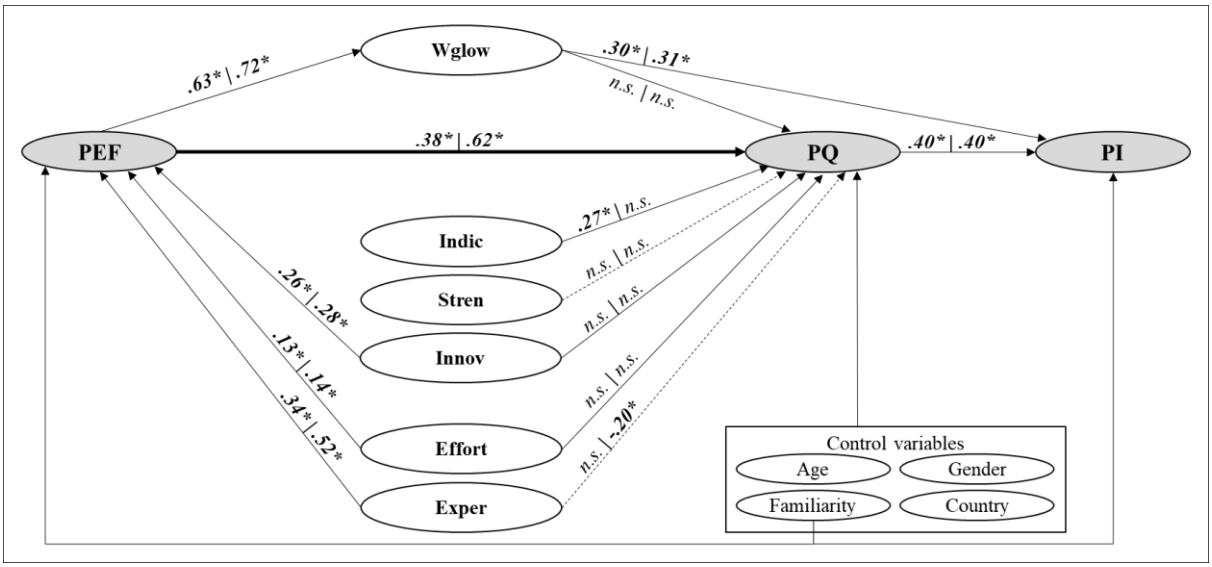


Figure 18: Visualization of SEM results for Sweater.
*Notes: Values before the slash show beta coefficients from the model with PEF_uni, and values after the slash beta coefficients from the model with PEF_dim. Results for control variables are not visualized here. *p < 0.05*

5.4.3. Comparison between all models

After the individual assessment of the analyzed models, results are compared between the models, focusing on the explained variances and the implications of the empirical results on the hypotheses. Analyzing the explained variances of the endogenous constructs shows an overall range between 31% for PEF in the Sweater condition (using PEF_uni) and 64% for Wglow in the Chocolate condition (using PEF_dim). Table 37 shows all R^2 values for the PLS SEM analyses. According to Hair et al. (2021), R^2 values of 0.50 are considered moderate and values of 0.25 are considered weak (values of 0.75 are considered substantial), although it also depends on the model and the respective complexity. Most R^2 in the SEM analyses in this thesis can be considered moderate. This is in line with expectations, in light of the many other known

influence factors towards PQ and PI, e.g., price, ambiguity tolerance or product knowledge (Singhal et al., 2019).

Table 37: Summary of explained variances of endogenous constructs across PLS SEM.

	Chocolate		Sweater	
Endogenous construct	PEF_uni	PEF_dim	PEF_uni	PEF_dim
PEF	0.353	0.467	0.307	0.549
Wglow	0.497	0.635	0.394	0.516
PQ	0.468	0.446	0.392	0.432
PI	0.350	0.349	0.436	0.436

Next, the overarching picture of path significances and thus support or lack of support for the hypotheses under investigation is summarized. Table 38 shows the results, indicating the β values for the different models (all shown values are significant at the 5% level, while n.s. indicates that the effect is not significant). When specifically comparing the models using PEF_uni versus PEF_dim, it can be seen that they are almost identical in terms of which hypothesized effects are significant. There are two exceptions where the results for PEF_uni and PEF_dim differ: First, for Chocolate, Stren influences PQ significantly when using PEF_dim, while it has no impact when using PEF_uni; Second, for Sweater, Indic influences PQ significantly only when using PEF_uni but not for PEF_dim. Beyond assessing which effects are significant, the β values are also of interest to assess if the strength of the relationship differs between PEF measures. It can be seen that β values are very similar or higher when using PEF_dim, and also the explained variances of the endogenous constructs are very similar or higher for PEF_dim. This indicates that the higher-order, multidimensional measure of PEF may be just as fitting or even better fitting in this context than a simple unidimensional measure as commonly used.

Table 38: Summary of support towards hypotheses across PLS SEM.
Notes: Values written in bold are statistically significant (p < 0.05); n.s. = not significant.

Hypothesis & Path			Chocolate		Sweater	
			PEF_uni (β)	PEF_dim (β)	PEF_uni (β)	PEF_dim (β)
1a	PEF	→ PQ	**0.356**	**0.352**	**0.367**	**0.620**
1b	PQ	→ PI	**0.320**	**0.321**	**0.400**	**0.399**
2a	PEF	→ Wglow	**0.706**	**0.798**	**0.629**	**0.719**
2b	Wglow	→ PQ	n.s.	n.s.	n.s.	n.s.
2c	Wglow	→ PI	**0.295**	**0.294**	**0.304**	**0.305**
3a	Indic	→ PQ	n.s.	n.s.	**0.269**	n.s.
4a	Stren	→ PQ	n.s.	**-0.122**	n.s.	n.s.
5a	Innov	→ PEF	**0.164**	**0.186**	**0.258**	**0.277**
5b	Innov	→ PQ	n.s.	n.s.	n.s.	n.s.
6a	Effort	→ PEF	n.s.	n.s.	**0.132**	**0.135**
6b	Effort	→ PQ	**0.118**	**0.129**	n.s.	n.s.
7a	Exper	→ PEF	**0.448**	**0.526**	**0.338**	**0.517**
7b	Exper	→ PQ	**0.204**	**0.191**	n.s.	**-0.195**
n/a	Age	→ PQ	**-0.129**	**-0.141**	**-0.097**	**-0.121**
n/a	Male	→ PQ	n.s.	n.s.	n.s.	n.s.
n/a	Germany	→ PQ	**0.090**	n.s.	n.s.	n.s.
n/a	Familiar	→ PEF	**0.109**	**0.129**	n.s.	n.s.
n/a	Familiar	→ PQ	n.s.	n.s.	n.s.	n.s.
n/a	Familiar	→ PI	n.s.	**0.143**	**0.208**	**0.207**

Summarizing the empirical findings in term of support for the hypotheses as shown in Table 38, it can be seen that some hypotheses are fully supported across all models, namely hypotheses 1a, 1b, 2a, 2c, 5a, and 7a, while others differ between products or PEF measures (hypotheses 3a, 4a 6a, 6b, and 7b). A few hypotheses need to be rejected based on the empirical results of this study, these are hypotheses 2b, 3b, 4b and 5b. Note that based on the previous regression analyses, hypotheses 3b and 4b (which described the interaction effects between Indic and PEF, as well as Stren and PEF) have not been included in the SEM analyses. However, to ensure that no premature interpretation was made, an additional PLS SEM analysis based on a model with both products jointly and PEF_uni as well as including the interaction effects was conducted. Results from this additional analysis support the irrelevance of the interaction effects: Both interaction effects were not significant, no path significances changed, and values of any other path coefficient changed only minimally when including the interaction effects.

5.5. Discussion of quantitative research

The work described in the last chapter extends the current knowledge on PEF, PQ and their interplay and correlations for multiple consumer products. The results will be interpreted in the next section and then the theoretical and practical implications will be described in the following sections. Afterwards, the limitations of the research will be discussed.

5.5.1. Interpretation of results

After having presented the statistical results in the last sections, this section is intended to interpret the key results before the following sections then describe the theoretical and practical implications. This quantitative analysis shed light onto several underlying questions, specifically relevance and impact of specific constructs within the context of PEF and PQ, different effects within the model for products of different industries and different effects of a uni- or multidimensional measure of PEF. Minimum requirements for sample sizes were met in all conditions, indicating that the statistical power was sufficient to detect the expected effects and that the results are robust. At the same time, the minimum sample sizes as assessed via G*Power were also not excessively exceeded (which might lead to statistical but not practical significant results) and effect sizes underlined the practical relevance of the effects. Reliability and validity indicators showed good fit for most measures indicating that the constructs were measured consistently and accurately, further supporting the robustness of the findings.

A direct and positive impact of PEF on PQ has been found consistently across products and PEF measures. Specifically for the Sweater, the relationship between PEF and PQ is much stronger when using PEF_dim as compared to any other model. This finding emphasizes the relevance of Dur not only for PEF but also for PQ – since PEF_dim for the Sweater contains Dur as one of the dimensions. The role of Dur for PQ has also been reported by other authors (Hazen et al., 2017). At the same time, PEF_dim contains Nat as dimension for both the Sweater and the Chocolate, and since there is virtually no difference in the path coefficients between PEF_uni and PEF_dim on PQ for Chocolate, Nat seems not to play a similarly big role for PQ of Chocolate. Given that Chocolate can be considered a hedonic product or also vice product (Van Doorn & Verhoef, 2011), future research could investigate the role of Nat for different kinds of products, e.g., fruit or vegetables. Taking a step back and considering the sample used for the quantitative research at hand, the rather large share of participants with low income could be another explanation for the emphasis on durability for PQ more than Nat. It could be the case that being durable is more important to this group and thus they put more emphasis on it in the quality assessment process.

The quantitative analyses provide once more evidence for the positive impact of PQ on PI. This has often been reported in literature (e.g., Das et al., 2020; Dodds et al., 1991) and again highlights the value of understanding und shaping consumers' quality perceptions. The values of the standardized path coefficients between PQ and PI are slightly higher for Sweater than Chocolate ($\beta_{PQ\text{-}PI}$ for Sweater: 0.40; $\beta_{PQ\text{-}PI}$ for Chocolate: 0.32), indicating that PQ plays a bigger role for the Sweater. This could be explained by the Sweater being a durable product and not a consumable like Chocolate and / or the Sweater being more expensive than the bar of Chocolate: Consumers might weigh PQ for PI more if they are spending more money and intending to use the product for a longer time.

The effect of PEF on Wglow is positive and strong across all conditions, showing that PEF evokes positive emotions in consumers. The effect of Wlog on PI is positive and virtually the same across all conditions. This in turn highlights the relevance of emotions on PI. Comparing the values of the path coefficients between PQ to PI and Wglow to PI, Wglow has only a slightly lower impact on PI than PQ does ($\Delta\beta = 0.03 – 0.10$, depending on the scenario). Overall, positive emotions can play a key role for PI and potentially purchase behavior on top of the impact of PQ. This effect seems not to be dependent on the type of product (Chocolate versus Sweater), meaning that alluding to consumers' positive emotions could be a marketing strategy feasible for a broad range of products.

Innov shows a consistent yet small to medium positive impact on PEF ($0.03 \leq f^2 \leq 0.13$). The effect size is bigger for Sweater than Chocolate, indicating potential differences of the role of product related factors like Innov. No specific information about the product's innovativeness was given in either product description, thus consumers seem to infer Innov from eco-friendly enhancements which in turn strengthens PEF. The role of more radical innovation remains yet to be further explored. At the same time, Innov shows no direct impact on PQ in the quantitative model at hand, which could be explained by the rather incremental innovativeness and by the indirect influence via PEF.

The remaining two product related constructs, Indic and Stren seem not to play a major role for PQ. Specifically, Indic has a small to medium positive effect on PQ only for Sweater with PEF_uni. However, Indic may still play a role for PQ in other scenarios, just not as explicitly and consciously measured variable: For example, future research could split consumers into groups depending on the presence or their level of Indic and then assess the direct impact of specific PEF dimension (e.g., Dur or Nat) on PQ. For Stren, only a small effect on PQ is found, and only for Chocolate with PEF_dim. This is an interesting finding for two

main reasons: First, Luchs et al. (2010) expected the strength-dependency to be focused on physical strength but not on a strong taste or smell. The negative effect of Stren on Chocolate however could suggest that beliefs about physical strength and effectiveness may be transferred by consumers towards other areas like taste. Future research could investigate this potential effect. Second, the lacking effect of Stren for the Sweater could indicate either that strength does not play a role for such a product or that consumers' conscious answers are biased. Mai et al. (2019) have already shown that implicit and explicitly expressed Stren can vary. Overall, it remains subject to future research to further investigate the relevance of Stren in terms of implicit beliefs as well as for different products where strength may play a more essential role (e.g., cleaning products or electronics like vacuum cleaners).

An unexpected result was the negative impact of Exper on PQ (in the model considering PEF_dim) for the Sweater. However, the effect size is small ($f^2 = 0.03$) and Exper in this scenario does positively impact PEF which in turn positively impacts PQ. Apart from that, Exper shows the expected positive impact on PQ for the Chocolate and on PEF for both products. For both products, the effect of Exper on PEF is stronger in the context of PEF_dim than in the context of PEF_uni, potentially indicating that Exper could be linked in consumers' minds more to specific PEF dimensions than a high-level universal PEF. Overall, the impact of Exper on PEF is medium to strong ($0.11 \leq f^2 \leq 0.40$), emphasizing the high relevance of perceptions towards producer attributes for product level PEF. Taking this a step further, future research could even investigate if the relationships between Exper and PEF could be mediated by (decreasing) greenwashing perceptions.

Effort shows differing results for Chocolate and Sweater: For Chocolate, it has a small, yet significant positive impact on PQ but not on PEF, while for Sweater it has a small, yet significant positive impact on PEF but not on PQ. This indicates that there could be different underlying mechanisms for how Effort is considered for specific products. However, the results for Effort need to be interpreted in light of its limitations regarding reliability and convergent validity and should be re-investigated in future research.

Regarding control variables, the findings are in line with previous research that found younger people to believe in higher PQ (Abbey et al., 2015; Van Doorn & Verhoef, 2011). However, the effect size of age is considered small in all scenarios that have been analyzed ($0.02 \leq f^2 \leq 0.04$). Gender is not a significant predictor of PQ as indicated by the PLS SEM results in this thesis. However, the initial regressions focused only on control variables showed the relevance of gender on PQ, in a way that men show lower PQ. It means that gender may

play a role depending on which other constructs are considered in the overall model. This is fully in line with previous research finding either being male as unfavorable for PQ (Tong & Su, 2018; Van Doorn & Verhoef, 2011) or having no impact on it (Ye et al., 2022). Looking at product familiarity, there is no effect in any scenario on PQ, but small effects ($0.02 \leq f^2 \leq 0.08$) on PEF and PI. Similar interlinkages of familiarity and PI have been found in other research (Jeon et al., 2024), and also for green claim skepticism (Kassie & Rhee, 2023). Even though the effects are small, they hint towards practical implications of increasing product familiarity.

Given the large diversity of countries of residence that are "Other" than Germany, the mostly non-significant (or minor / without practical relevance) effects ($f^2 \leq 0.02$) of country of longest residence are not surprising – especially since previous research also showed mixed results in terms of geographical differences (Kaczorowska et al., 2021; Osburg et al., 2024). The geographical coverage is considered a limitation of this study and further outlined in section 5.5.4. Further control variables were not included in the PLS SEM due to the identification step via MLR. Control variables for which data had been collected but regression analyses showed no relevant impact within the analyzed model are the following: Having children, household size, living area, highest education, household income.

5.5.2. Theoretical implications

After having described and interpreted the findings from the quantitative research at hand, this section outlines its implications for research before the following section will describe the practical implications. Theoretical implications can be grouped into five areas:

First, the quantitative study emphasizes the relevance of PEF for PQ as shown by the consistently positive effect of PEF on PQ. This extends existing research and suggests changing perceptions of consumers towards seeing PEFPs in a more positive way. Future research should further assess and potentially confirm this pattern.

Second, this study confirms that PEF does not only increase PQ but also influences consumers on an emotional level, i.e., raising Wglow. In fact, the impact of PEF on Wglow showed consistently the strongest effect sizes within the whole model, highlighting the relevance of this relationship. Beyond the direct effects of PEF on PQ and Wglow, both these constructs increase PI, thus suggesting an indirect positive influence of PEF on PI via PQ and Wglow. Future research should thus further investigate the role of consumers' (positive) emotions within the context of eco-friendly behaviors.

Third, this study provides evidence for the relevance of the proposed multidimensional PEF conceptualization. Confirmatory factor analyses showed initial support for the proposed

dimensionality. Parallel-test reliability was indicated by correlating the unidimensional and multidimensional measures. Initial support for the conceptualization was also given by assessment of internal consistency reliability, convergent validity, discriminant validity and nomological validity via SEM considering the higher-order PEF within a network of other theoretical constructs. Thus, future research should further investigate and validate these initial findings, including empirically assessing the proposed set of dimensions, identifying relevant dimensions for a broader set of products and developing or refining appropriate measures.

Fourth, the specific product under investigation matters and could change research results (as already initially indicated in the systematic literature review). In the quantitative study at hand, a Chocolate bar and a Sweater had been chosen (based on previous research, examples from the interviews, and a specific pretest). While some results are the same for both products (e.g., the direct relationships from PEF to PQ and from PQ to PI) others differ (e.g., the impact of theoretical constructs like Indic or Stren). Future research should thus pay specific attention to the products used in research and systematically extend the body of knowledge for specific products and product types.

Fifth, the quantitative results suggest that socio-demographic factors play a limited role in shaping PQ within the studied context. Despite the initial inclusion of a broad set of control variables, their minor impact in MLR and then SEM analyses indicates that other factors, such as consumer, product, or perceiver oriented variables, are more influential. This suggests that researchers may want to carefully consider which socio-demographic variables to collect or not to collect, e.g., due to the sensitive nature and potential irritation of study participants.

5.5.3. Practical implications

The quantitative study within this thesis provides implications for companies, which are outlined in this section. First, this study provides guidance for companies that aim to make their eco-friendly products more attractive and increase PI of consumers. The positive effect of familiarity on PEF and PI could be leveraged by companies, for example with strategies like free samples to increase product familiarity and thus overcome barriers regarding consumers' perceptions and intentions. Additionally, companies can make use of the impact of anticipated positive emotions (Wglow) on PI and not only advertise the product's functional features but rather address the emotional layer and effects that the product will have for consumers. Second, the positive impact of PEF on PQ and ultimately on PI should motivate companies to invest in EF of their products. Third, companies may historically base their strategies for targeted marketing on socio-demographic consumer information. However, this study suggests that

there may only be limited value in targeting marketing efforts based on factors like gender, or even income. Thus, a more successful strategy could be to invest resources rather into shaping consumer perceptions, e.g., about the eco-expertise of the own brand. Fourth, when trying to shape PEF beyond addressing factors like the own brand image, companies can leverage the specific PEF dimensions and provide specific information about the relevant dimensions to be more concrete and avoid consumer confusion.

5.5.4. Limitations of conducted quantitative research

The quantitative research described in this chapter has some specific limitations, which can be grouped into four areas which are explained in more detail in this section: Measure reliability / validity constraints, lack of behavioral measure and limited geographical coverage / systematic assessment.

First, three measures used in this study (Effort, Nat, and Res) showed low values for convergent validity (AVE below 0.5). The low convergent validity of Nat and Res only influences the dimensional measure of PEF (PEF_dim), thus it does not cause concerns for calculation and interpretation of the models operationalized via PEF_uni. For Effort, the reliability indicators were just above the recommended thresholds, which means that the results in terms of Effort should be considered with a grain of salt and reinvestigated by future research. Overall, the low convergent validity (and reliability) could be explained by the following reasoning: 1) The original measure items were translated which might cause slight differences in meaning despite careful translation. 2) The measures for Nat and Res were initially developed in the context of food and thus may need to be adjusted to better fit other product types (i.e., the Sweater in this study). In addition, even though the measure used for Nat was published about 10 years ago (Pula et al., 2014), its original underlying measure was developed almost 30 years ago (Steptoe et al., 1995). The underlying construct may have changed over time and thus further adjustment of the measure could be considered going forward. 3) The measure for Effort in this study was phrased to capture general perceptions of producers rather than the perception of one concrete producer – so that for example heterogeneity in perceptions or different interpretation (e.g., considering all consumer goods producers versus only sweater producers) could cause low reliability and convergent validity. To address this specifically, future research could investigate Effort specifically for a concrete producer. Additionally, the measure for Effort contained a reverse item, which might have confused participants so that future research could apply a different measure without reverse items and then compare results.

An additional limitation derived from the constructs' reliability and validity assessments, is found for Indic and Rec as self-developed measures were used for these

constructs. While the reliability and validity indicators were satisfactory for the measure of Indic, a proper measure validation would be useful. Given the novelty of this construct in research and the limited findings in the quantitative research at hand, another qualitative assessment might even be useful initially to assess the applicability in a different context. For Rec, a potential revision and even questioning if it should be split into two separate constructs might be beneficial. The reliability and validity indicators for this construct were not satisfactory, requiring more attention than a measure validation. Another alternative might be to consider Rec as formative construct and re-assessing it from that perspective.

Third, in the conducted study, PI was used as proxy for behavior, which is not necessarily true as shown by the intention-behavior gap documented in literature. Since using actual behavior would be more costly and complicated, and the model under investigation introduced novel constructs in the context of PEF, it was deemed fitting to conduct the initial investigations using PI. The results of the SEM analyses seem promosing so that investigating impact on actual behavior may be useful as a follow-up in future research.

Fourth, in terms of geographical reach, the current study did mostly focus on Germany and while including consumers from other countries, it did not systematically assess differences between any other countries or regions. Results indicated some differences between German consumers and consumers from other countries, so that future research should further investigate the model in different geographical contexts and systematically compare more granular differences.

5.6. Summary of quantitative research

The quantitative research part of this thesis has evaluated the proposed relationships between PEF, PQ, and PI, as well as the influence of several proposed constructs (i.e. Exper, Innov, Wglow) and (socio-demographic) control variables influencing these relationships. A convenience sample of 732 observations (after data cleansing) was recruited via three main channels: 1) From the network of the author of this thesis; 2) Via Prolific (for a small participation fee); 3) Via SurveyCircle and SurveySwap. The online survey software (Unipark) randomized the allocation of participants to either the Chocolate (367 observations) or the Sweater (365 observations) condition. The demographic composition of the final sample shows a rather young sample (Mean = 32 years, Median = 28 years, SD = 11 years), with 79% having no own children, 56% having reported a net household income of < 50k € in the past year and 45% of the sample identifying as male. After confirmation of the measures and suitability for further analyses via CFA, MLR analyses were used to ensure parsimony of the effects under

investigation and to confirm the suitability of the sample. Finally, structural equation modelling (PLS SEM) was used to investigate the proposed models. Four different conditions were analyzed separately: For two different products (Chocolate and Sweater) and for two different measures of PEF (unidimensional and multidimensional). The positive influence of PEF on PQ and of PQ on PI has been shown in all models as has the positive effect of PEF on Wglow and of Wglow on PI. In addition, Innov consistently shows a positive impact on PEF, but no impact on PQ. The influence of the other proposed constructs and control variables on PEF, PQ and PI varies between conditions. These findings propose several contributions to research and practice, for example providing further evidence for the relevance of PEF towards PQ and PI, highlighting the potential of emotional effects (i.e., Wglow), providing initial support for a novel (multidimensional) conceptualization of PEF, and differentiating research findings for specific products (implying the need to differentiate practical strategies between different products). The quantitative study has some limitations, specifically regarding constraints in measure reliability and validity, lacking assessment of actual consumer behavior and lacking systematic assessment of geographical differences.

6. Overall discussion

This section intends to connect the insights and implications of all previous parts of this thesis. It starts by integrating the results of the systematic literature review, the qualitative study, and the quantitative assessment. Subsequently, theoretical and practical implications are discussed and a description of the limitations spanning the whole thesis concludes this work.

6.1. Integration of results

This thesis has aimed to answer the two key questions of how PEFPs are defined and conceptualized, and what the relationships between PEF and PQ are. Doing so, it has first identified working definitions, then systematically reviewed the existing literature, followed by an interview-based qualitative study and finally a quantitative assessment of the developed model. This mixed methods approach has allowed to develop a well-rounded view and generate both depth and breadth to answer the key research questions. Answers to the research questions are provided in the following.

1) How are PEFPs defined and conceptualized?

The developed definition for PEFP is the following: A PEFP is a product that consumers perceive to have a more positive impact or a lower, not long-lasting, or preferably no harmful impact on the environment (non-living and living things, including animals and humans) across its whole life-cycle (before, during and after use) compared to alternatives that fulfill the same consumer need.

Beyond the definition, this thesis has proposed a novel conceptualization of PEF as higher-order multidimensional construct based on dimensions like naturalness, social friendliness, animal friendliness, climate friendliness, resource efficiency, durability, repairability or reuse- and recyclability (depending on the specific product and context). Quantitative assessment has provided first evidence for the applicability of (at least some of) these dimensions for Chocolate and Sweater. Findings from this thesis also suggest that the specific understanding of and expectations towards PEF depend on the product type / industry, on perceiver characteristics (e.g., knowledge about or experiences with PEFPs) and situational context (e.g., social influences of other people or product marketing).

2) What are the relationships between PEF and PQ?

This thesis has developed a nuanced view of the relationships between PEF and PQ. Based on the literature review, a multi-faceted picture was created both for the direct impact of PEF on PQ, and for additional theoretical constructs impacting these relationships. It was found

that the direct impact can be negative, neutral or positive depending on the specific research setting, including e.g., when the research was conducted, which specific products were used for assessment, or which contextual information was available to research participants. In the quantitative research within this thesis, the direct impact of PEF on PQ is clearly and consistently positive. As far as additional theoretical constructs are concerned, this thesis has clustered them into consumer / perceiver oriented, situational / product oriented and producer (or more broadly firm) oriented. Specific constructs within all these categories impact the relationships between PEF and PQ in their own ways, once more highlighting the complexity of this topic. Some constructs such as ambiguity tolerance of the perceiver (Hazen et al., 2012) or more concrete product information available (Jäger & Weber, 2020) show a positive impact, while others such as perceived disease threat (Lu & Kwan, 2023) or environmental skepticism (Jäger & Weber, 2020; Royne et al., 2012) show a negative impact and again others such as perceived producer eco-expertise or strength beliefs show different impact depending on the product assessed.

6.2. Overarching theoretical implications

Beyond the specific implications as presented in the individual section of this thesis, there are overarching theoretical implications. Six major implications for research are summarized and described in this section.

First, this thesis has summarized and systemized the existing research and findings about relationships between PEF and PQ. In doing so, it has taken a cross-industry perspective and created an extensive overview as well as framework for future research to build upon. Over the course of the thesis, the empirical investigations have already started to add onto this extensive overview.

Second, this thesis has reconciled the scientific and consumer definitions of PEF, by analyzing existing definitions in scientific literature and refining the initially taken working definition within this thesis via consumer interviews. It has shown the relevance of not only environmental aspects but also considerations towards social impacts and animal welfare.

Third, in addition to reconciling PEF definitions, this thesis has also proposed a novel conceptualization of PEF as higher-order multidimensional construct. This conceptualization developed in this thesis was informed by interview data and started to be validated through existing literature, ensuring both theoretical grounding and practical relevance. The quantitative study provided initial evidence for the usefulness and applicability of the multidimensional

higher-order conceptualization. Future research could further assess conceptualization and especially validate empirical measures.

Fourth, this thesis has identified new constructs and relationships within the PEF – PQ context. An initial set of additional constructs with potential effects on PEF – PQ relationships were proposed after systematically reviewing prior studies, as outlined in the description of research gaps in section 3.5.3. Furthermore, the qualitative study identified specific constructs, such as Exper or Wglow, while the quantitative study assessed the statistical relevance of these constructs for two different products (Chocolate and Sweater) and different PEF measures.

Fifth, this thesis highlights the importance of product type and industry in research about PEF and PQ. This became evident both through a systematic assessment of existing literature and the comparison of two distinct products in the quantitative study. Specifically, this thesis has differentiated findings between product types / industries and highlighted the importance of considering specific products. Consequently, the transferability of findings may be limited to products of the same type or within the same industry. Researchers should thus carefully consider the product specifics within previous research, critically assess implications on their own hypotheses and design, and provide enough details of their own research so that others can use findings in an informed way.

Sixth, the importance of context and specifically the point in time emerged as recurring theme for PEF – PQ relationships from the systematic literature review and the qualitative interviews. This implies the need to reassess past studies across different (temporal) contexts and to consider the circumstances and shifting paradigms in future research.

6.3. Overarching practical implications

This section summarizes the practical implications based on the analyses in this thesis. The implications are presented as recommendations addressed towards policy makers and companies. Which recommendation(s) policy makers or companies may want to apply remains their own choice.

Policy makers should:

- Ensure that consumers are properly educated and have the capabilities to assess PEF of everyday consumer products
- Ensure that consumers not only have the capabilities to assess PEF but also have the relevant information at hand or easily available, e.g., by designing appropriate and clear labelling schemes, and setting specific guidelines for companies

- Consider the environmental, social and animal friendliness aspects of PEF within regulations, campaigns or similar

Companies should:

- Not only focus on the factual basis and functional aspects of products but also address the emotional layer (without manipulating consumers), e.g., by advertising how PEFPs will make consumers feel
- Actively shape their own brand image, e.g., towards high perceived eco-expertise – for which they may potentially consider a multiband strategy since a) consumers seem to associate younger and smaller brands with fewer greenwashing behaviors, and b) a holistic eco-friendly image and product range of one specific brand seem to be favored over singular eco-friendly advertised product lines
- Consider carefully which retailers to collaborate with because retailer image may influence both greenwashing perceptions / PEF as well as PQ of the products offered
- Target marketing efforts to specific consumers based on PEF-related variables like product familiarity or knowledge and not (only) based on generic socio-demographic variables
- Try to actively increase familiarity of consumers with their eco-friendly products, e.g., via free samples or free return offers
- Sensitize marketing and communications teams to shape product advertisements towards increasing PEF and PQ of specific products, e.g., by providing concrete information about EF, providing more than one cue for EF and combining EF information with quality-related information
- Consider the dimensionality of PEF, e.g., to provide specific information about the relevant dimensions of a product to achieve being more concrete
- Consider the specific nature of each product and remember that successful strategies for one product may not necessarily translate to another product
- Consider the environmental, social and animal friendliness aspects associated with PEF by consumers, e.g., in product design and marketing campaigns, to provide an overarching and integrated view towards PEF
- Encourage internal cross-functional cooperation, e.g., between marketing and product development to ensure a coherent view and optimization of the product's story

- Consider and optimize not only the core product but also the effect of its packaging, e.g., regarding packaging material, sizing and design, including white space
- Collaborate across the whole value chain, with NGOs and potentially even competitors, to increase customer knowledge and overall interest in PEFPs, while also positively impacting their own brand image

6.4. Overarching limitations

There are some overarching limitations that span across the different parts of this thesis, which are described in this section. Specific limitations for the systematic literature review, the qualitative study and the quantitative study have been discussed in the respective chapters.

Overall, even though this thesis has applied different methods (e.g., interview-based GT and online survey-based structural equation modelling), its focus is still rather limited to specific methods. The same can be said to a large extent about the literature within the systematic literature review. Thus, expanding the methods used to assess relationships between PEF and PQ could further add evidence and increase the knowledge base within this area. Potential methods to be used in future research could include IAT, fsQCA or other methods. For example, an IAT could be used to investigate the subconscious impact of specific dimensions of PEF (e.g., perceived naturalness and perceived social friendliness) on the overarching PEF as well as on PQ. It could also investigate the impact of subconscious beliefs of constructs like perceived product innovativeness or perceived producer eco-effort on PQ in the context of PEF. Another method, fsQCA, could be leveraged to identify specific configurations of perceiver, product and producer oriented attributes that lead to higher or lower PQ. FsQCA could also be applied to better understand which combinations of product attributes (e.g. brand, price, origin) and perceiver variables (e.g. environmental concern, anticipated warm glow) can increase PEF and / or PQ.

Additionally, while the quantitative research in this thesis has differentiated analyses and results by product, the systematic literature review and qualitative assessment have not. This aligns with previous research that focuses on specific products. However, as demonstrated by the differences between Chocolate and Sweater, as well as findings from prior studies (Petersen & Brockhaus, 2017; Van Doorn & Verhoef, 2011; Weisstein et al., 2024), results should be interpreted with careful consideration of the specific product. There are various ways to classify products, such as consumable vs. durable, low vs. high involvement, hedonic vs. utilitarian, simple vs. sophisticated, or low vs. high strength-related. Systematically clustering

and analyzing these classifications in future research could enhance the transferability of findings and improve their practical applicability.

The samples used in this thesis may be another factor potentially limiting the generalizability of the results. Both the qualitative and quantitative research focused mainly on Germany, although other countries were included to a smaller extent. Further research is needed to examine whether the findings hold across different regions and cultures. In addition, even though the interview participants were carefully selected and theoretical saturation was reached, a study with different participants might have created different results. Similarly, the sample for the quantitative study was not representative, and although the investigated products were chosen carefully, using different products could lead to different outcomes. Finally, even the articles in the systematic literature review needed to be selected, and alternative search terms or inclusion / exclusion criteria could lead to different results. Despite clear selection criteria and repeated literature analyses, some bias in article selection cannot be entirely ruled out.

Finally, researchers and practitioners need to be aware that all findings and implications presented in this thesis are related to a specific moment in time and might change over time. For example, during the completion of this thesis in spring 2025, a major political shift occurred in the USA – particularly reducing the importance of sustainability and eco-friendliness from the perspective of the government. As such political and other (social, economic, etc.) changes can have an impact on the perceptions of individuals, it seems worthwhile for future research to re-assess the questions and findings presented in this thesis.

REFERENCES

Aaker, D. A. (1991). *Managing Brand Equity: Capitalizing on the Value of a Brand Name* (Illustrated Edition). The Free Press.

Abbey, J. D., Kleber, R., Souza, G. C., & Voigt, G. (2017). The Role of Perceived Quality Risk in Pricing Remanufactured Products. *Production & Operations Management*, *26*(1), 100–115. https://doi.org/10.1111/poms.12628

Abbey, J. D., & Meloy, M. G. (2017). Attention by design: Using attention checks to detect inattentive respondents and improve data quality. *Journal of Operations Management*, *53–56*(1), 63–70. https://doi.org/10.1016/j.jom.2017.06.001

Abbey, J. D., Meloy, M. G., Guide, V. D. R., & Atalay, S. (2015). Remanufactured Products in Closed-Loop Supply Chains for Consumer Goods. *Production & Operations Management*, *24*(3), 488–503. https://doi.org/10.1111/poms.12238

Abdulla, H., Abbey, J. D., Atalay, A. S., & Meloy, M. G. (2024). Show, don't tell: Education and physical exposure effects in remanufactured product markets. *Journal of Operations Management*, *70*(2), 243–256. https://doi.org/10.1002/joom.1248

Achabou, M. A., & Dekhili, S. (2013). Luxury and sustainable development: Is there a match? *Journal of Business Research*, *66*(10), 1896–1903. https://doi.org/10.1016/j.jbusres.2013.02.011

Achabou, M. A., Dekhili, S., & Codini, A. P. (2020). Consumer preferences towards animal-friendly fashion products: An application to the Italian market. *Journal of Consumer Marketing*, *37*(6), 661–673. https://doi.org/10.1108/JCM-10-2018-2908

Aćimović, S., Mijušković, V., & Rajić, V. (2020). The impact of reverse logistics onto green supply chain competitiveness evidence from Serbian consumers. *International Journal of Retail & Distribution Management*, *48*(9), 1003–1021. https://doi.org/10.1108/IJRDM-04-2019-0142

Agrawal, V. V., Atasu, A., & van Ittersum, K. (2015). Remanufacturing, Third-Party Competition, and Consumers' Perceived Value of New Products. *Management Science*, *61*(1), 60–72. https://doi.org/10.1287/mnsc.2014.2099

Agrawal, V. V., Ferguson, M., Toktay, L. B., & Thomas, V. M. (2012). Is Leasing Greener Than Selling? *Management Science*, *58*(3), 523–533. https://doi.org/10.2139/ssrn.1857535

Ahluwalia, R., Unnava, H. R., & Burnkrant, R. E. (2001). The Moderating Role of Commitment on the Spillover Effect of Marketing Communications. *Journal of Marketing Research*, *38*(4), 458–470. https://doi.org/10.1509/jmkr.38.4.458.18903

Ahmad, W., & Zhang, Q. (2020). Green purchase intention: Effects of electronic service quality and customer green psychology. *Journal of Cleaner Production*, *267*, 122053. https://doi.org/10.1016/j.jclepro.2020.122053

Akaichi, F., Glenk, K., & Revoredo-Giha, C. (2019). Could animal welfare claims and nutritional information boost the demand for organic meat? Evidence from non-hypothetical experimental auctions. *Journal of Cleaner Production*, *207*, 961–970. https://doi.org/10.1016/j.jclepro.2018.10.064

Alexis Dinno. (2014). *dunn.test: Dunn's Test of Multiple Comparisons Using Rank Sums* (p. 1.3.6) [Dataset]. https://doi.org/10.32614/CRAN.package.dunn.test

American Psychological Association. (2018a). *APA Dictionary of Psychology*. Relationship. https://dictionary.apa.org/

American Psychological Association. (2018b). *APA Dictionary of Psychology*. Measure. https://dictionary.apa.org/

Ammann, J., Mack, G., El Benni, N., Jin, S., Newell-Price, P., Tindale, S., Hunter, E., Vicario-Modroño, V., Gallardo-Cobos, R., Sánchez-Zamora, P., Miškolci, S., & Frewer, L. J. (2024). Consumers across five European countries prioritise animal welfare above environmental sustainability when buying meat and dairy products. *Food Quality and Preference*, *117*, 105179. https://doi.org/10.1016/j.foodqual.2024.105179

Anagnostou, A., Ingenbleek, P. T. M., & van, T. H. C. M. (2015). Sustainability labelling as a challenge to legitimacy: Spillover effects of organic Fairtrade coffee on consumer perceptions of mainstream products and retailers. *Journal of Consumer Marketing*, *32*(6), 422–431. https://doi.org/10.1108/JCM-11-2014-1213

Andreoni, J. (1990). Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving. *The Economic Journal*, *100*(401), 464–477. https://doi.org/10.2307/2234133

Antonetti, P., & Maklan, S. (2014a). Exploring Postconsumption Guilt and Pride in the Context of Sustainability. *Psychology & Marketing*, *31*(9), 717–735. https://doi.org/10.1002/mar.20730

Antonetti, P., & Maklan, S. (2014b). Feelings that Make a Difference: How Guilt and Pride Convince Consumers of the Effectiveness of Sustainable Consumption Choices. *Journal of Business Ethics*, *124*(1), 117–134. https://doi.org/10.1007/s10551-013-1841-9

Antonetti, P., Manika, D., & Katsikeas, C. (2019). Why consumer animosity reduces product quality perceptions: The role of extreme emotions in international crises. *International Business Review*, *28*(4), 739–753. https://doi.org/10.1016/j.ibusrev.2019.02.003

Atasu, A., Sarvary, M., & Van Wassenhove, L. N. (2008). Remanufacturing as a Marketing Strategy. *Management Science*, *54*(10), 1731–1746. https://doi.org/10.1287/mnsc.1080.0893

Atkinson, L., & Rosenthal, S. (2014). Signaling the Green Sell: The Influence of Eco-Label Source, Argument Specificity, and Product Involvement on Consumer Trust. *Journal of Advertising*, *43*(1), 33–45. https://doi.org/10.1080/00913367.2013.834803

Baek, E., Choo, H. J., & Lee, S. H. (Mark). (2018). Using warmth as the visual design of a store: Intimacy, relational needs, and approach intentions. *Journal of Business Research*, *88*, 91–101. https://doi.org/10.1016/j.jbusres.2018.03.013

Bagozzi, R. P. (2010). Structural equation models are modelling tools with many ambiguities: Comments acknowledging the need for caution and humility in their use. *Journal of Consumer Psychology*, *20*(2), 208–214. https://doi.org/10.1016/j.jcps.2010.03.001

Bai, S., Zhang, X., Han, C., & Yu, D. (2023). Research on the Influence Mechanism of Organic Food Attributes on Customer Trust. *Sustainability*, *15*(8), Article 8. https://doi.org/10.3390/su15086733

Baker, M. A., Shin, J. T., & Kim, Y. W. (2016). An Exploration and Investigation of Edible Insect Consumption: The Impacts of Image and Description on Risk Perceptions and

Purchase Intent. *Psychology & Marketing*, *33*(2), 94–112. https://doi.org/10.1002/mar.20847

Bangsa, A. B., & Schlegelmilch, B. B. (2020). Linking sustainable product attributes and consumer decision-making: Insights from a systematic review. *Journal of Cleaner Production*, *245*, 118902. https://doi.org/10.1016/j.jclepro.2019.118902

Bao, Y., Bao, Y., & Sheng, S. (2011). Motivating purchase of private brands: Effects of store image, product signatureness, and quality variation. *Journal of Business Research*, *64*(2), 220–226. https://doi.org/10.1016/j.jbusres.2010.02.007

Barbarossa, C., & De Pelsmacker, P. (2016). Positive and Negative Antecedents of Purchasing Eco-friendly Products: A Comparison Between Green and Non-green Consumers. *Journal of Business Ethics*, *134*(2), 229–247. https://doi.org/10.1007/s10551-014-2425-z

Bartels, R., & Goodhew, J. (1981). The Robustness of the Durbin-Watson Test. *The Review of Economics and Statistics*, *63*(1), 136–139. https://doi.org/10.2307/1924228

Bartlett, M. S. (1950). Tests of Significance in Factor Analysis. *British Journal of Statistical Psychology*, *3*(2), 77–85. https://doi.org/10.1111/j.2044-8317.1950.tb00285.x

Basfirinci, C., & Cilingir Uk, Z. (2020). Does Country of Origin Matter for Chocolate? Ethnocentrism, Involvement, and Perceived Risk for Turkish University Students. *Journal of Food Products Marketing*, *26*(2), 144–184. https://doi.org/10.1080/10454446.2020.1740128

Beamon, B. M. (1999). Designing the green supply chain. *Logistics Information Management*, *12*(4), 332–342. https://doi.org/10.1108/09576059910284159

Bechwati, N. N., & Xia, L. (2003). Do Computers Sweat? The Impact of Perceived Effort of Online Decision Aids on Consumers' Satisfaction With the Decision Process. *Journal of Consumer Psychology*, *13*(1–2), 139–148. https://doi.org/10.1207/S15327663JCP13-1&2_12

Becker, J.-M., Ringle, C. M., & Sarstedt, M. (2018). ESTIMATING MODERATING EFFECTS IN PLS-SEM AND PLSc-SEM: INTERACTION TERM GENERATION*DATA TREATMENT. *Journal of Applied Structural Equation Modeling*, *2*(2), 1–21. https://doi.org/10.47263/JASEM.2(2)01

Becker-Olsen, K., & Potucek, S. (2013). Greenwashing. In S. O. Idowu, N. Capaldi, L. Zu, & A. D. Gupta (Eds.), *Encyclopedia of Corporate Social Responsibility*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-28036-8

Berinsky, A. J., Margolis, M. F., & Sances, M. W. (2014). Separating the Shirkers from the Workers? Making Sure Respondents Pay Attention on Self-Administered Surveys. *American Journal of Political Science*, *58*(3), 739–753. https://doi.org/10.1111/ajps.12081

Bernstein, D. A., & Nash, P. W. (2008). *Essentials of psychology* (4th ed). Houghton Mifflin Co.

Besharat, A. (2010). How co-branding versus brand extensions drive consumers' evaluations of new products: A brand equity approach. *Industrial Marketing Management*, *39*(8), 1240–1249. https://doi.org/10.1016/j.indmarman.2010.02.021

Bezençon, V., Girardin, F., & Lunardo, R. (2020). When does an ethical attribute matter for product evaluation? The role of warm-glow feelings for low-rated products. *Psychology & Marketing*, *37*(11), 1571–1585. https://doi.org/10.1002/mar.21403

Bhander, G. S., Hauschild, M., & McAloone, T. (2003). Implementing life cycle assessment in product development. *Environmental Progress*, *22*(4), 255–267. https://doi.org/10.1002/ep.670220414

Bhardwaj, A. K., Garg, A., Ram, S., Gajpal, Y., & Zheng, C. (2020). Research Trends in Green Product for Environment: A Bibliometric Perspective. *International Journal of Environmental Research and Public Health*, *17*(22), 8469. https://doi.org/10.3390/ijerph17228469

Bhatt, S., Lee, J., Deutsch, J., Ayaz, H., Fulton, B., & Suri, R. (2018). From food waste to value-added surplus products (VASP): Consumer acceptance of a novel food product category. *Journal of Consumer Behaviour*, *17*(1), 57–63. https://doi.org/10.1002/cb.1689

Bigliardi, B., Campisi, D., Ferraro, G., Filippelli, S., Galati, F., & Petroni, A. (2020). The Intention to Purchase Recycled Products: Towards an Integrative Theoretical Framework. *Sustainability*, *12*(22), Article 22. https://doi.org/10.3390/su12229739

Binninger, A.-S. (2017). Perception of Naturalness of Food Packaging and Its Role in Consumer Product Evaluation. *Journal of Food Products Marketing*, *23*(3), 251–266. Business Source Ultimate.

Blanco-Penedo, I., García-Gudiño, J., Angón, E., Perea, J. M., Escribano, A. J., & Font-i-Furnols, M. (2021). Exploring Sustainable Food Choices Factors and Purchasing Behavior in the Sustainable Development Goals Era in Spain. *Sustainability*, *13*(13), Article 13. https://doi.org/10.3390/su13137397

Blasch, J., Filippini, M., & Kumar, N. (2019). Boundedly rational consumers, energy and investment literacy, and the display of information on household appliances. *Resource & Energy Economics*, *56*, 39–58. bsu.

Blasi, S., Brigato, L., & Sedita, S. R. (2020). Eco-friendliness and fashion perceptual attributes of fashion brands: An analysis of consumers' perceptions based on twitter data mining. *Journal of Cleaner Production*, *244*, 118701. https://doi.org/10.1016/j.jclepro.2019.118701

Blokhuis, H., Jones, B., Geers, R., Miele, M., & Veissier, I. (2003). Measuring and Monitoring Animal Welfare: Transparency in the Food Product Quality Chain. *Animal Welfare 12 (2003) 4*, *12*. https://doi.org/10.1017/S096272860002604X

Bodur, H. O., Gao, T., & Grohmann, B. (2014). The Ethical Attribute Stigma: Understanding When Ethical Attributes Improve Consumer Responses to Product Evaluations. *Journal of Business Ethics*, *122*(1), 167–177. https://doi.org/10.1007/S10551-013-1764-5

Bodur, H. O., Tofighi, M., & Grohmann, B. (2016). When Should Private Label Brands Endorse Ethical Attributes? *Journal of Retailing*, *92*(2), 204–217. https://doi.org/10.1016/j.jretai.2015.11.001

Boesen, S., Bey, N., & Niero, M. (2019). Environmental sustainability of liquid food packaging: Is there a gap between Danish consumers' perception and learnings from

life cycle assessment? *Journal of Cleaner Production*, *210*, 1193–1206. https://doi.org/10.1016/j.jclepro.2018.11.055

Boisvert, J., & Khan, M. S. (2022). Toward a better understanding of the main antecedents and outcomes of consumer-based perceived product innovativeness. *Journal of Strategic Marketing*, *30*(3), 296–319. https://doi.org/10.1080/0965254X.2020.1807589

Boks, C., & Stevels, A. (2007). Essential perspectives for design for environment. Experiences from the electronics industry. *International Journal of Production Research*, *45*(18–19), 4021–4039. https://doi.org/10.1080/00207540701439909

Bom, S., Jorge, J., Ribeiro, H. M., & Marto, J. (2019). A step forward on sustainability in the cosmetics industry: A review. *Journal of Cleaner Production*, *225*, 270–290. https://doi.org/10.1016/j.jclepro.2019.03.255

Borin, N., Cerf, D. C., & Krishnan, R. (2011). Consumer effects of environmental impact in product labeling. *Journal of Consumer Marketing*, *28*(1), 76–86. https://doi.org/10.1108/07363761111101976

Borsboom, D., Mellenbergh, G. J., & Van Heerden, J. (2003). The theoretical status of latent variables. *Psychological Review*, *110*(2), 203–219. https://doi.org/10.1037/0033-295X.110.2.203

Borsboom, D., Mellenbergh, G. J., & Van Heerden, J. (2004). The Concept of Validity. *Psychological Review*, *111*(4), 1061–1071. https://doi.org/10.1037/0033-295X.111.4.1061

Bosmans, A. (2006). Scents and Sensibility: When Do (In)Congruent Ambient Scents Influence Product Evaluations? *Journal of Marketing*, *70*(3), 32–43. https://doi.org/10.1509/jmkg.70.3.032

Box, G. E. P., & Cox, D. R. (1964). An Analysis of Transformations. *Journal of the Royal Statistical Society. Series B (Methodological)*, *26*(2), 211–252.

Boz, Z., Korhonen, V., & Koelsch Sand, C. (2020). Consumer Considerations for the Implementation of Sustainable Packaging: A Review. *Sustainability*, *12*(6), Article 6. https://doi.org/10.3390/su12062192

Brach, S., Walsh, G., & Shaw, D. (2018). Sustainable consumption and third-party certification labels: Consumers' perceptions and reactions. *European Management Journal*, *36*(2), 254–265. https://doi.org/10.1016/j.emj.2017.03.005

Bracquené, E., Peeters, J., Alfieri, F., Sanfélix, J., Duflou, J., Dewulf, W., & Cordella, M. (2021). Analysis of evaluation systems for product repairability: A case study for washing machines. *Journal of Cleaner Production*, *281*, 125122. https://doi.org/10.1016/j.jclepro.2020.125122

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

Breusch, T. S., & Pagan, A. R. (1979). A Simple Test for Heteroscedasticity and Random Coefficient Variation. *Econometrica*, *47*(5), 1287–1294. https://doi.org/10.2307/1911963

Brockman, B. K., & Morgan, R. M. (2003). The Role of Existing Knowledge in New Product Innovativeness and Performance. *Decision Sciences*, *34*(2), 385–419. https://doi.org/10.1111/1540-5915.02326

Bronnmann, J., Stoeven, M. T., Quaas, M., & Asche, F. (2021). Measuring Motivations for Choosing Ecolabeled Seafood: Environmental Concerns and Warm Glow. *Land Economics*, *97*(3), 641–654. https://doi.org/10.3368/le.97.3.641

Brough, A. R., Wilkie, J. E. B., Ma, J., Isaac, M. S., & Gal, D. (2016). Is Eco-Friendly Unmanly? The Green-Feminine Stereotype and Its Effect on Sustainable Consumption. *Journal of Consumer Research*, *43*(4), 567–582. https://doi.org/10.1093/jcr/ucw044

Brown, T. (2023). Confirmatory Factor Analysis. In R. H. Hoyle (Ed.), *Handbook of Structural Equation Modeling* (2nd ed.). Guilford Publications.

Bryła, P. (2019). Regional Ethnocentrism on the Food Market as a Pattern of Sustainable Consumption. *Sustainability*, *11*(22), Article 22. https://doi.org/10.3390/su11226408

Bühner, M. (2011). *Einführung in die Test- und Fragebogenkonstruktion—Das Standardwerk der empirischen Psychologie in neuer Bearbeitung (Pearson Studium—Psychologie) von Markus Bühner* (3rd ed.). Person Studium.

Bühner, M., & Ziegler, M. (2009). *Statistik für Psychologen und Sozialwissenschaftler*. Pearson Deutschland GmbH.

Bürgin, D., & Wilken, R. (2021). Increasing consumers' purchase intentions toward fair-trade products through partitioned pricing. *Journal of Business Ethics*, *181*, 1015–1040. https://doi.org/10.1007/s10551-021-04938-6

Burke, P. F., Eckert, C., & Davis, S. (2014). Segmenting consumers' reasons for and against ethical consumption. *European Journal of Marketing*, *48*(11/12), 2237–2261. https://doi.org/10.1108/EJM-06-2013-0294

Calantone, R. J., Chan, K., & Cui, A. S. (2006). Decomposing Product Innovativeness and Its Effects on New Product Success. *Journal of Product Innovation Management*, *23*(5), 408–421. https://doi.org/10.1111/j.1540-5885.2006.00213.x

Camacho-Otero, J., Boks, C., & Pettersen, I. (2018). Consumption in the Circular Economy: A Literature Review. *Sustainability*, *10*(8), 2758. https://doi.org/10.3390/su10082758

Campbell, B., Khachatryan, H., Behe, B., Dennis, J., & Hall, C. (2015). Consumer Perceptions of Eco-friendly and Sustainable Terms. *Agricultural and Resource Economics Review*, *44*, 21–34. https://doi.org/10.1017/S1068280500004603

Can, B. A. (2023). Turkish Consumers' Perceptions of Organic Milk and the Factors Affecting Consumption: The Case of Kocaeli, Türkiye. *Sustainability*, *15*(13), Article 13. https://doi.org/10.3390/su151310044

Canty, A. J., & Ripley, B. (2024). *boot: Bootstrap R (S-Plus) Functions* (Version R package version 1.3-30.) [Computer software].

Chakraborty, S., & Newton, A. C. (2011). Climate change, plant diseases and food security: An overview. *Plant Pathology*, *60*(1), 2–14. https://doi.org/10.1111/j.1365-3059.2010.02411.x

Challagalla, G., Murtha, B. R., & Jaworski, B. (2014). Marketing Doctrine: A Principles-Based Approach to Guiding Marketing Decision Making in Firms. *Journal of Marketing*, *78*(4), 4–20. https://doi.org/10.1509/jm.12.0314

Chandra Pant, S., Saxena, R., Gupta, N. K., Yadav, H., Kumar AD, S., & Kumar Pant, D. (2024). The organic odyssey: Navigating the influence of attitude on purchase intent, mediated by perceived value, quality, and price in India. *Journal of Retailing and Consumer Services*, *79*, 103801. https://doi.org/10.1016/j.jretconser.2024.103801

Chandy, R. K., & Tellis, G. J. (2000). The Incumbent's Curse? Incumbency, Size, and Radical Product Innovation. *Journal of Marketing*, *64*(3), 1–17. https://doi.org/10.1509/jmkg.64.3.1.18033

Chang, M.-Y., & Chen, H.-S. (2022). Consumer Attitudes and Purchase Intentions in Relation to Animal Welfare-Friendly Products: Evidence from Taiwan. *Nutrients*, *14*(21), 4571. https://doi.org/10.3390/nu14214571

Charmaz, K. (2006). *Constructing grounded theory*. SAGE Publications.

Charmaz, K. (2014). *Constructing Grounded Theory* (2nd ed.). SAGE Publications.

Charmaz, K., & Belgrave, L. L. (2015). Grounded Theory. In *The Blackwell Encyclopedia of Sociology*. American Cancer Society. https://doi.org/10.1002/9781405165518.wbeosg070.pub2

Charmaz, K., & Thornberg, R. (2021). The pursuit of quality in grounded theory. *Qualitative Research in Psychology*, *18*(3), 305–327. https://doi.org/10.1080/14780887.2020.1780357

Chase, K. S., & Murtha, B. (2019). Selling to Barricaded Buyers. *Journal of Marketing*, *83*(6), 2–20. https://doi.org/10.1177/0022242919874778

Chebat, J.-C., & Michon, R. (2003). Impact of ambient odors on mall shoppers' emotions, cognition, and spending: A test of competitive causal theories. *Journal of Business Research*, *56*(7), 529–539. https://doi.org/10.1016/S0148-2963(01)00247-8

Chen, X., Gao, Z., Swisher, M., House, L., & Zhao, X. (2018). Eco-labeling in the Fresh Produce Market: Not All Environmentally Friendly Labels Are Equally Valued. *Ecological Economics*, *154*, 201–210. https://doi.org/10.1016/j.ecolecon.2018.07.014

Chen, Y., Ghosh, M., Liu, Y., & Zhao, L. (2019). Media Coverage of Climate Change and Sustainable Product Consumption: Evidence from the Hybrid Vehicle Market. *Journal of Marketing Research*, *56*(6), 995–1011. https://doi.org/10.1177/0022243719865898

Chen, Y.-S., & Chang, C.-H. (2013). Towards green trust: The influences of green perceived quality, green perceived risk, and green satisfaction. *Management Decision*, *51*(1), 63–82. https://doi.org/10.1108/00251741311291319

Chen, Y.-S., Lin, C.-Y., & Weng, C.-S. (2015). The Influence of Environmental Friendliness on Green Trust: The Mediation Effects of Green Satisfaction and Green Perceived Quality. *Sustainability*, *7*(8), Article 8. https://doi.org/10.3390/su70810135

Cheung, G. W., Cooper-Thomas, H. D., Lau, R. S., & Wang, L. C. (2024). Reporting reliability, convergent and discriminant validity with structural equation modeling: A review and best-practice recommendations. *Asia Pacific Journal of Management*, *41*(2), 745–783. https://doi.org/10.1007/s10490-023-09871-y

Cho, E., & Kim, S. (2015). Cronbach's Coefficient Alpha: Well Known but Poorly Understood. *Organizational Research Methods*, *18*(2), 207–230. https://doi.org/10.1177/1094428114555994

Cho, Y.-N., & Berry, C. (2019). Understanding the effects of retailer- and manufacturer-provided sustainability labels on product evaluations and purchase-related outcomes.

Journal of Business Research, *100*, 73–85. https://doi.org/10.1016/j.jbusres.2019.03.019

Choi, B.-C., Shin, H.-S., Lee, S.-Y., & Hur, T. (2006). Life Cycle Assessment of a Personal Computer and its Effective Recycling Rate (7 pp). *The International Journal of Life Cycle Assessment*, *11*(2), 122–128. https://doi.org/10.1065/lca2004.12.196

Chueh, T. Y., & Kao, D. T. (2004). The moderating effects of consumer perception to the impacts of country-of-design on perceived quality. *Journal of American Academy of Business*, *4*(1/2), 70–74. https://doi.org/10.1080/00913367.2001.10673652

Chung, J., & Monroe, G. S. (2003). Exploring Social Desirability Bias. *Journal of Business Ethics*, *44*(4), 291–302.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed). L. Erlbaum Associates.

Commission of the European Communities. (2001). *Green paper on integrated product policy*. Commission of the European Communities. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52001DC0068&rid=1

Cooper, D. R., & Gutowski, T. G. (2017). The Environmental Impacts of Reuse: A Review. *Journal of Industrial Ecology*, *21*(1), 38–56. https://doi.org/10.1111/jiec.12388

Cooper, D. R., & Schindler, P. S. (2014). *Business research methods* (12th ed.). McGraw-Hill/Irwin.

Corbin, J. M., & Strauss, A. (1990). Grounded theory research: Procedures, canons, and evaluative criteria. *Qualitative Sociology*, *13*(1), 3–21. https://doi.org/10.1007/BF00988593

Corley, K., & Gioia, D. A. (2011). Building theory about theory building: What constitutes a theoretical contribution? *The Academy of Management Review*, *36*(1), 12–32. https://doi.org/10.5465/amr.2009.0486

Creswell, J. W. (2008). *Research design: Qualitative, quantitative, and mixed methods approaches* (2. ed., [Nachdr.]). SAGE Publications.

Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches* (3rd ed). SAGE Publications.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*(3), 297–334. https://doi.org/10.1007/BF02310555

Dangelico, R. M., & Pontrandolfo, P. (2010). From green product definitions and classifications to the Green Option Matrix. *Journal of Cleaner Production*, *18*(16), 1608–1628. https://doi.org/10.1016/j.jclepro.2010.07.007

Dangelico, R. M., & Vocalelli, D. (2017). "Green Marketing": An analysis of definitions, strategy steps, and tools through a systematic review of the literature. *Journal of Cleaner Production*, *165*, 1263–1279. https://doi.org/10.1016/j.jclepro.2017.07.184

Darnall, N., Ji, H., & Vázquez-Brust, D. A. (2018). Third-Party Certification, Sponsorship, and Consumers' Ecolabel Use. *Journal of Business Ethics*, *150*(4), 953–969. https://doi.org/10.1007/s10551-016-3138-2

Das, G., Peloza, J., Varshneya, G., & Green, T. (2020). When do consumers value ethical attributes? The role of perceived quality in gift-giving. *European Journal of Marketing*, *55*(1), 315–335. https://doi.org/10.1108/EJM-02-2019-0202

Das Guru, R. R., & Paulssen, M. (2020). Customers' experienced product quality: Scale development and validation. *European Journal of Marketing*, *54*(4), 645–670. https://doi.org/10.1108/EJM-03-2018-0156

Davies, I. A., & Gutsche, S. (2016). Consumer motivations for mainstream "ethical" consumption. *European Journal of Marketing*, *50*(7/8), 1326–1347. https://doi.org/10.1108/EJM-11-2015-0795

Davis, D. F., Golicic, S. L., & Boerstler, C. N. (2011). Benefits and challenges of conducting multiple methods research in marketing. *Journal of the Academy of Marketing Science*, *39*(3), 467–479. https://doi.org/10.1007/s11747-010-0204-7

Davison, A. C., & Hinkley, D. V. (1997). *Bootstrap Methods and their Applications.* Cambridge University Press.

De Fano, D., Schena, R., & Russo, A. (2022). Empowering plastic recycling: Empirical investigation on the influence of social media on consumer behavior. *Resources, Conservation and Recycling*, *182*, 106269. https://doi.org/10.1016/j.resconrec.2022.106269

De Freitas Netto, S. V., Sobral, M. F. F., Ribeiro, A. R. B., & Soares, G. R. da L. (2020). Concepts and forms of greenwashing: A systematic review. *Environmental Sciences Europe*, *32*(1), 19. https://doi.org/10.1186/s12302-020-0300-3

De Medeiros, J. F., Ribeiro, J. L. D., & Cortimiglia, M. N. (2014). Success factors for environmentally sustainable product innovation: A systematic literature review. *Journal of Cleaner Production*, *65*, 76–86. https://doi.org/10.1016/j.jclepro.2013.08.035

De Medeiros, J. F., Ribeiro, J. L. D., & Cortimiglia, M. N. (2016). Influence of perceived value on purchasing decisions of green products in Brazil. *Journal of Cleaner Production*, *110*, 158–169. https://doi.org/10.1016/j.jclepro.2015.07.100

De Oña, J., & De Oña, R. (2015). Quality of Service in Public Transport Based on Customer Satisfaction Surveys: A Review and Assessment of Methodological Approaches. *Transportation Science*, *49*(3), 605–622. https://doi.org/10.1287/trsc.2014.0544

De Pelsmacker, P., & Janssens, W. (2007). A model for fair trade buying behaviour: The role of perceived quantity and quality of information and of product-specific attitudes. *Journal of Business Ethics*, *75*(4), 361–380. https://doi.org/10.1007/s10551-006-9259-2

DeepL SE. (n.d.). *DeepL Translate* [Computer software]. https://www.deepl.com/de/translator

Dekhili, S., & Achabou, M. A. (2013). Price Fairness in the Case of Green Products: Enterprises' Policies and Consumers' Perceptions. *Business Strategy and the Environment*, *22*(8), 547–560. https://doi.org/10.1002/bse.1763

Dekhili, S., & Achabou, M. A. (2015). The Influence of the Country-of-Origin Ecological Image on Ecolabelled Product Evaluation: An Experimental Approach to the Case of the European Ecolabel. *Journal of Business Ethics*, *131*(1), 89–106. https://doi.org/10.1007/s10551-014-2261-1

Delmas, M. A., & Burbano, V. C. (2011). The Drivers of Greenwashing. *California Management Review*, *54*(1), 64–87. https://doi.org/10.1525/cmr.2011.54.1.64

Delmas, M. A., & Gergaud, O. (2021). Sustainable practices and product quality: Is there value in eco-label certification? The case of wine. *Ecological Economics*, *183*, 106953. https://doi.org/10.1016/j.ecolecon.2021.106953

Deo, K., & Prasad, A. (2024). Factors influencing green energy consumer behaviour in Australia. *Journal of Cleaner Production*, *460*, 142609. https://doi.org/10.1016/j.jclepro.2024.142609

Derdowski, L. A., Grahn, Å. H., Hansen, H., & Skeiseid, H. (2020). The New Ecological Paradigm, Pro-Environmental Behaviour, and the Moderating Effects of Locus of Control and Self-Construal. *Sustainability*, *12*(18), Article 18. https://doi.org/10.3390/su12187728

Diamantopoulos, A., Sarstedt, M., Fuchs, C., Wilczynski, P., & Kaiser, S. (2012). Guidelines for choosing between multi-item and single-item scales for construct measurement: A predictive validity perspective. *Journal of the Academy of Marketing Science*, *40*(3), 434–449. https://doi.org/10.1007/s11747-011-0300-3

Diamantopoulos, A., & Siguaw, J. A. (2006). Formative Versus Reflective Indicators in Organizational Measure Development: A Comparison and Empirical Illustration. *British Journal of Management*, *17*(4), 263–282. https://doi.org/10.1111/j.1467-8551.2006.00500.x

Diamantopoulos, A., & Winklhofer, H. M. (2001). Index Construction with Formative Indicators: An Alternative to Scale Development. *Journal of Marketing Research*, *38*(2), 269–277. https://doi.org/10.1509/jmkr.38.2.269.18845

Dixon, D., & Mikolon, S. (2021). Cents of self: How and when self-signals influence consumer value derived from choices of green products. *International Journal of Research in Marketing*, *38*(2), 365–386. https://doi.org/10.1016/j.ijresmar.2020.08.002

Dodds, W. B., & Monroe, K. B. (1985). The Effect of Brand and Price Information on Subjective Product Evaluations. *ACR North American Advances*, *NA-12*. https://www.acrwebsite.org/volumes/6364/volumes/v12/NA-12/full

Dodds, W. B., Monroe, K. B., & Grewal, D. (1991). Effects of Price, Brand, and Store Information on Buyers' Product Evaluations. *Journal of Marketing Research*, *28*(3), 307–319. JSTOR. https://doi.org/10.2307/3172866

Drost, E. A. (2011). Validity and Reliability in Social Science Research. *Education Research and Perspectives*, *38*(1), 105–123.

D'Souza, C., Taghian, M., & Lamb, P. (2006). An empirical study on the influence of environmental labels on consumers. *Corporate Communications: An International Journal*, *11*(2), 162–173. https://doi.org/10.1108/13563280610661697

Du, H., & Bentler, P. M. (2022). 40-Year Old Unbiased Distribution Free Estimator Reliably Improves SEM Statistics for Nonnormal Data. *Structural Equation Modeling: A Multidisciplinary Journal*, *29*(6), 872–887. https://doi.org/10.1080/10705511.2022.2063870

Duan, Y., & Aloysius, J. A. (2019). Supply chain transparency and willingness-to-pay for refurbished products. *The International Journal of Logistics Management*, *30*(3), 797–820. https://doi.org/10.1108/IJLM-01-2019-0025

Dunn, T. J., Baguley, T., & Brunsden, V. (2014). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. *British Journal of Psychology*, *105*(3), 399–412. https://doi.org/10.1111/bjop.12046

Durbin, J., & Watson, G. S. (1950). Testing for Serial Correlation in Least Squares Regression: I. *Biometrika*, *37*(3/4), 409–428. https://doi.org/10.2307/2332391

Durbin, J., & Watson, G. S. (1951). Testing for Serial Correlation in Least Squares Regression. II. *Biometrika*, *38*(1/2), 159–177. https://doi.org/10.2307/2332325

Durbin, J., & Watson, G. S. (1971). Testing for Serial Correlation in Least Squares Regression. III. *Biometrika*, *58*(1), 1–19. https://doi.org/10.2307/2334313

Durif, F., Boivin, C., & Julien, C. (2010). In Search of a Green Product Definition. *Innovative Marketing*, *6*(1), 25–33.

Dutta-Powell, R., Rhee, J. J., & Wodak, S. (2024). Two interventions for mitigating the harms of greenwashing on consumer perceptions. *Business Strategy and the Environment*, *33*(2), 882–903. https://doi.org/10.1002/bse.3520

Dziuban, C. D., & Shirkey, E. C. (1974). When is a correlation matrix appropriate for factor analysis? Some decision rules. *Psychological Bulletin*, *81*(6), 358–361. https://doi.org/10.1037/h0036316

Edwards, J. R. (2001). Multidimensional Constructs in Organizational Behavior Research: An Integrative Analytical Framework. *Organizational Research Methods*, *4*(2), 144–192. https://doi.org/10.1177/109442810142004

Eid, M., Gollwitzer, M., & Schmitt, M. (2011). *Statistik und Forschungsmethoden: Mit Online-Materialien* (2nd ed.). Beltz.

Elgaaïed-Gambier, L. (2016). Who Buys Overpackaged Grocery Products and Why? Understanding Consumers' Reactions to Overpackaging in the Food Sector. *Journal of Business Ethics*, *135*(4), 683–698. https://doi.org/10.1007/s10551-014-2491-2

ElHaffar, G., Durif, F., & Dubé, L. (2020). Towards closing the attitude-intention-behavior gap in green consumption: A narrative review of the literature and an overview of future research directions. *Journal of Cleaner Production*, *275*, 122556. https://doi.org/10.1016/j.jclepro.2020.122556

Elliott, G. R., & Cameron, R. C. (1994). Consumer Perception of Product Quality and the Country-of-Origin Effect. *Journal of International Marketing*, *2*(2), 49–62. https://doi.org/10.1177/1069031X9400200204

Epskamp, S. (2022). *semPlot: Path Diagrams and Visual Analysis of Various SEM Packages' Output* (Version R package version 1.1.6) [Computer software]. https://CRAN.R-project.org/package=semPlot

European Commission. (2016, November 23). *Climate change consequences* [Text]. Climate Action - European Commission. https://ec.europa.eu/clima/change/consequences_en

European Commission. (2020). *Special Eurobarometer 501. Summary. Attitudes of Europeans towards the Environment.* European Union. https://europa.eu/eurobarometer/surveys/detail/2257

Fachbach, I., Lechner, G., & Reimann, M. (2022). Drivers of the consumers' intention to use repair services, repair networks and to self-repair. *Journal of Cleaner Production*, *346*, 130969. https://doi.org/10.1016/j.jclepro.2022.130969

Farías, P. (2020). Promoting the Absence of Pesticides through Product Labels: The Role of Showing a Specific Description of the Harmful Effects, Environmental Attitude, and Familiarity with Pesticides. *Sustainability*, *12*(21), Article 21. https://doi.org/10.3390/su12218912

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Fernando, A. G., Sivakumaran, B., & Suganthi, L. (2018). Comparison of perceived acquisition value sought by online second-hand and new goods shoppers. *European Journal of Marketing*, *52*(7/8), 1412–1438. https://doi.org/10.1108/EJM-01-2017-0048

Fernando, A. G., Suganthi, L., & Sivakumaran, B. (2014). If You Blog, Will They Follow? Using Online Media to Set the Agenda for Consumer Concerns on "Greenwashed" Environmental Claims. *Journal of Advertising*, *43*(2), 167–180. https://doi.org/10.1080/00913367.2013.841088

Feucht, Y., & Zander, K. (2018). Consumers' preferences for carbon labels and the underlying reasoning. A mixed methods approach in 6 European countries. *Journal of Cleaner Production*, *178*, 740–748. https://doi.org/10.1016/j.jclepro.2017.12.236

Fischer, D. G., & Fick, C. (1993). Measuring Social Desirability: Short Forms of the Marlowe-Crowne Social Desirability Scale. *Educational and Psychological Measurement*, *53*(2), 417–424. https://doi.org/10.1177/0013164493053002011

Flint, D. J., Woodruff, R. B., & Gardial, S. F. (2002). Exploring the Phenomenon of Customers' Desired Value Change in a Business-to-Business Context. *Journal of Marketing*, *66*(4), 102–117. https://doi.org/10.1509/jmkg.66.4.102.18517

Fogt Jacobsen, L., Pedersen, S., & Thøgersen, J. (2022). Drivers of and barriers to consumers' plastic packaging waste avoidance and recycling – A systematic literature review. *Waste Management*, *141*, 63–78. https://doi.org/10.1016/j.wasman.2022.01.021

Forbes, S. L., Cohen, D. A., Cullen, R., Wratten, S. D., & Fountain, J. (2009). Consumer attitudes regarding environmentally sustainable wine: An exploratory study of the New Zealand marketplace. *Journal of Cleaner Production*, *17*(13), 1195–1199. https://doi.org/10.1016/j.jclepro.2009.04.008

Fornell, C., & Larcker, D. F. (1981). Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. *Journal of Marketing Research*, *18*(1), 39–50. https://doi.org/10.2307/3151312

Fox, J., & Weisberg, S. (2019). *An R Companion to Applied Regression* (Third). Sage. https://socialsciences.mcmaster.ca/jfox/Books/Companion/

Franco, A., Malhotra, N., & Simonovits, G. (2014). Publication bias in the social sciences: Unlocking the file drawer. *Science*, *345*(6203), 1502–1505. https://doi.org/10.1126/science.1255484

Frank, P., & Brock, C. (2019). "Green cannibalism" or an "organic inside job"? Empirical insights into the rivalry of ethical grocery types. *Psychology & Marketing*, *36*(6), 597–617. Business Source Ultimate.

Frota Neto, J. Q., Bloemhof, J., & Corbett, C. (2016). Market prices of remanufactured, used and new items: Evidence from eBay. *International Journal of Production Economics*, *171*(Part 3), 371–380. https://doi.org/10.1016/j.ijpe.2015.02.006

Fuchs, C., Schreier, M., & van Osselaer, S. M. J. (2015). The handmade effect: What's love got to do with it? *Journal of Marketing*, *79*(2), 98–110. https://doi.org/10.1509/jm.14.0018

Furlow, N. E. (2010). Greenwashing in the New Millennium. *Journal of Applied Business and Economics*, *10*(6), 22–26.

Gagne, P., & Hancock, G. R. (2006). Measurement Model Quality, Sample Size, and Solution Propriety in Confirmatory Factor Models. *Multivariate Behavioral Research*, *41*(1), 65–83. https://doi.org/10.1207/s15327906mbr4101_5

Galesic, M. (2006). Dropouts on the Web: Effects of Interest and Burden Experienced During an Online Survey. *Journal of Official Statistics*, *22*(2), 313–328.

Gallarza, M. G., Gil-Saura, I., & Holbrook, M. B. (2011). The value of value: Further excursions on the meaning and role of customer value. *Journal of Consumer Behaviour*, *10*(4), 179–191. https://doi.org/10.1002/cb.328

Garvin, D. A. (1987, November 1). Competing on the Eight Dimensions of Quality. *Harvard Business Review*, *November 1987*, 101–109.

Gassler, B., von Meyer-Höfer, M., & Spiller, A. (2016). Exploring Consumers' Expectations of Sustainability in Mature and Emerging Markets. *Journal of Global Marketing*, *29*(2), 71–84. https://doi.org/10.1080/08911762.2015.1133869

Gaukroger, S., Schuster, J. A., & Sutton, J. (Eds.). (2000). *Descartes' natural philosophy*. Routledge.

Gaur, J., Amini, M., Banerjee, P., & Gupta, R. (2015). Drivers of consumer purchase intentions for remanufactured products: A study of Indian consumers relocated to the USA. *Qualitative Market Research: An International Journal*, *18*(1), 30–47. https://doi.org/10.1108/QMR-01-2014-0001

Geissdoerfer, M., Savaget, P., Bocken, N. M. P., & Hultink, E. J. (2017). The Circular Economy – A new sustainability paradigm? *Journal of Cleaner Production*, *143*, 757–768. https://doi.org/10.1016/j.jclepro.2016.12.048

Geldermann, J., Kolbe, L. M., Krause, A., Mai, C., Militz, H., Osburg, V.-S., Schoebel, A., Schumann, M., Toporowski, W., & Westphal, S. (2016). Improved resource efficiency and cascading utilisation of renewable materials. *Journal of Cleaner Production*, *110*, 1–8. https://doi.org/10.1016/j.jclepro.2015.09.092

German Environmental Agency. (2024). *Indicator: Environmentally friendly consumption*. Indicator: Environmentally Friendly Consumption; Umweltbundesamt. https://www.umweltbundesamt.de/en/data/environmental-indicators/indicator-environmentally-friendly-consumption

Gershoff, A. D., & Frels, J. K. (2015). What Makes It Green? The Role of Centrality of Green Attributes in Evaluations of the Greenness of Products. *Journal of Marketing*, *79*(1), 97–110. https://doi.org/10.1509/jm.13.0303

Gioia, D. A., Corley, K. G., & Hamilton, A. L. (2013). Seeking Qualitative Rigor in Inductive Research: Notes on the Gioia Methodology. *Organizational Research Methods*, *16*(1), 15–31. https://doi.org/10.1177/1094428112452151

Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Aldine.

Gleim, M., & Lawson, S. J. (2014). Spanning the gap: An examination of the factors leading to the green gap. *Journal of Consumer Marketing*, *31*(6/7), 503–514. https://doi.org/10.1108/JCM-05-2014-0988

Gleim, M., Smith, J., Andrews, D., & Cronin. Jr, J. J. (2013). Against the Green: A Multi-method Examination of the Barriers to Green Consumption. *Journal of Retailing*, *89*, 44–61. https://doi.org/10.1016/j.jretai.2012.10.001

Gobert, J., Allais, R., & Deroubaix, J.-F. (2021). Repair and reuse: Misalignments between stakeholders and possible users. *Journal of Cleaner Production*, *317*, 128454. https://doi.org/10.1016/j.jclepro.2021.128454

Golder, P. N., Mitra, D., & Moorman, C. (2012). What is Quality? An Integrative Framework of Processes and States. *Journal of Marketing*, *76*(4), 1–23. https://doi.org/10.1509/jm.09.0416

Goldstein, E. B., & Brockmole, J. R. (2017). *Sensation and Perception* (10th ed.). Cengage Learning.

Google. (n.d.). *Google Translate*. https://translate.google.com/?hl=de&sl=auto&tl=en&op=translate

Gorton, M., Tocco, B., Yeh, C.-H., & Hartmann, M. (2021). What determines consumers' use of eco-labels? Taking a close look at label trust. *Ecological Economics*, *189*, 107173. https://doi.org/10.1016/j.ecolecon.2021.107173

Gosselt, J. F., van Rompay, T., & Haske, L. (2019). Won't Get Fooled Again: The Effects of Internal and External CSR ECO-Labeling. *Journal of Business Ethics*, *155*(2), 413–424. https://doi.org/10.1007/s10551-017-3512-8

Goulding, C. (2002). *Grounded theory: A practical guide for management, business and market researchers*. SAGE Publications.

Graedel, T. E., Allenby, B. R., & Comriε, P. R. (1995). Matrix Approaches to Abridged Life Cycle Assessment. *Environmental Science & Technology*, *29*(3), 134A-139A. https://doi.org/10.1021/es00003a751

Greenwood, S. C., Walker, S., Baird, H. M., Parsons, R., Mehl, S., Webb, T. L., Slark, A. T., Ryan, A. J., & Rothman, R. H. (2021). Many Happy Returns: Combining insights from the environmental and behavioural sciences to understand what is required to make reusable packaging mainstream. *Sustainable Production and Consumption*, *27*, 1688–1702. https://doi.org/10.1016/j.spc.2021.03.022

Grene, M. (1974). *The Understanding of Nature* (Vol. 23). D. Reidel Publishing Company. https://doi.org/10.1007/978-94-010-2224-8

Grewal, D., Monroe, K. B., & Krishnan, R. (1998). The Effects of Price-Comparison Advertising on Buyers' Perceptions of Acquisition Value, Transaction Value, and Behavioral Intentions. *Journal of Marketing*, *62*(2), 46–59. https://doi.org/10.2307/1252160

Grinstein, A., & Nisan, U. (2009). Demarketing, Minorities, and National Attachment. *Journal of Marketing*, *73*(2), 105–122. https://doi.org/10.1509/jmkg.73.2.105

Groening, C., Sarkis, J., & Zhu, Q. (2018). Green marketing consumer-level theory review: A compendium of applied theories and further research directions. *Journal of Cleaner Production*, *172*, 1848–1866. https://doi.org/10.1016/j.jclepro.2017.12.002

Grolleau, G., Mzoughi, N., & Sutan, A. (2019). Does advertising the green benefits of products contribute to sustainable development goals? A quasi-experimental test of the dilution effect. *Business Strategy and the Environment*, *28*(5), 786–793. https://doi.org/10.1002/bse.2280

Grunert, K. G. (2002). Current issues in the understanding of consumer food choice. *Trends in Food Science & Technology*, *13*(8), 275–285. https://doi.org/10.1016/S0924-2244(02)00137-1

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

Gupta, R., & Sen, S. (2013). The effect of evolving resource synergy beliefs on the intentions–behavior discrepancy in ethical consumption. *Journal of Consumer Psychology*, *23*(1), 114–121. https://doi.org/10.1016/j.jcps.2012.07.004

Gutiérrez, S., Cortés, S., Ruedlinger, J., Ojeda, M. J., & Fresán, U. (2023). Geographic variability in the Chilean dietary carbon footprint: Major food contributors and associated factors. *Journal of Cleaner Production*, *427*, 139118. https://doi.org/10.1016/j.jclepro.2023.139118

Habel, J., Schons, L. M., Alavi, S., & Wieseke, J. (2016). Warm Glow or Extra Charge? The Ambivalent Effect of Corporate Social Responsibility Activities on Customers' Perceived Price Fairness. *Journal of Marketing*, *80*(1), 84–105. https://doi.org/10.1509/jm.14.0389

Hagtvedt, H., & Brasel, S. A. (2017). Color Saturation Increases Perceived Product Size. *Journal of Consumer Research*, *44*(2), 396–413. https://doi.org/10.1093/jcr/ucx039

Hair, J. F., Anderson, R., Babin, B., & Black, W. (2013). *Multivariate Data Analysis*. Pearson Deutschland. https://elibrary.pearson.de/book/99.150005/9781292035116

Hair, J. F., Black, W., Babin, B. J., & Anderson, D. R. (2018). Supplements. In *Multivariate Data Analysis* (8th ed.). Cengage Learning. https://mvstats.com/supplements/

Hair, J. F., Howard, M. C., & Nitzl, C. (2020). Assessing measurement model quality in PLS-SEM using confirmatory composite analysis. *Journal of Business Research*, *109*, 101–110. https://doi.org/10.1016/j.jbusres.2019.11.069

Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2022). *A primer on partial least squares structural equation modeling (PLS-SEM)*. Sage Publishing.

Hair, J. F., Hult, G. T. M., Ringle, C. M., Sarstedt, M., Danks, N. P., & Ray, S. (2021). *Partial Least Squares Structural Equation Modeling (PLS-SEM) Using R: A Workbook*. Springer International Publishing. https://doi.org/10.1007/978-3-030-80519-7

Hair, J. F., Hult, G. T. M., Ringle, C. M., Sarstedt, M., & Thiele, K. O. (2017). Mirror, mirror on the wall: A comparative evaluation of composite-based structural equation modeling methods. *Journal of the Academy of Marketing Science*, *45*(5), 616–632. https://doi.org/10.1007/s11747-017-0517-x

Hair, J. F., Ringle, C. M., & Sarstedt, M. (2011). Pls-Sem: Indeed a Silver Bullet. *Journal of Marketing Theory and Practice*, *19*(2), 139–151. https://doi.org/10.2753/MTP1069-6679190202

Hair, J. F., Sarstedt, M., Ringle, C. M., & Gudergan, S. P. (2024). *Advanced Issues in Partial Least Squares Structural Equation Modeling* (2nd edition). SAGE Publications.

Halog, A., & Manik, Y. (2011). Advancing Integrated Systems Modelling Framework for Life Cycle Sustainability Assessment. *Sustainability*, *3*(2), Article 2. https://doi.org/10.3390/su3020469

Harrell Jr, F. E. (2024). *Hmisc: Harrell Miscellaneous* (Version R package version 5.1-3) [Computer software]. https://CRAN.R-project.org/package=Hmisc

Hartikainen, H., Roininen, T., Katajajuuri, J.-M., & Pulkkinen, H. (2014). Finnish consumer perceptions of carbon footprints and carbon labelling of food products. *Journal of Cleaner Production*, *73*, 285–293. https://doi.org/10.1016/j.jclepro.2013.09.018

Hartmann, P., & Apaolaza-Ibáñez, V. (2012). Consumer attitude and purchase intention toward green energy brands: The roles of psychological benefits and environmental concern. *Journal of Business Research*, *65*(9), 1254–1263. https://doi.org/10.1016/j.jbusres.2011.11.001

Hartmann, P., Eisend, M., Apaolaza, V., & D'Souza, C. (2017). Warm glow vs. altruistic values: How important is intrinsic emotional reward in proenvironmental behavior? *Journal of Environmental Psychology*, *52*, 43–55. https://doi.org/10.1016/j.jenvp.2017.05.006

Hayes, A. F., & Coutts, J. J. (2020). Use Omega Rather than Cronbach's Alpha for Estimating Reliability. But…. *Communication Methods and Measures*, *14*(1), 1–24. https://doi.org/10.1080/19312458.2020.1718629

Hazen, B. T., Boone, C. A., Wang, Y., & Khor, K. S. (2017). Perceived quality of remanufactured products: Construct and measure development. *Journal of Cleaner Production*, *142*, 716–726. https://doi.org/10.1016/j.jclepro.2016.05.099

Hazen, B. T., Cegielski, C., & Hanna, J. B. (2011). Diffusion of green supply chain managementExamining perceived quality of green reverse logistics. *International Journal of Logistics Management*, *22*(3), 373–389. https://doi.org/10.1108/09574091111181372

Hazen, B. T., Overstreet, R. E., Jones-Farmer, L. A., & Field, H. S. (2012). The role of ambiguity tolerance in consumer perception of remanufactured products. *International Journal of Production Economics*, *135*(2), 781–790. https://doi.org/10.1016/j.ijpe.2011.10.011

Hecke, T. V. (2012). Power study of anova versus Kruskal-Wallis test. *Journal of Statistics and Management Systems*, *15*(2–3), 241–247. https://doi.org/10.1080/09720510.2012.10701623

Heene, M., Hilbert, S., Draxler, C., Ziegler, M., & Bühner, M. (2011). Masking misfit in confirmatory factor analysis by increasing unique variances: A cautionary note on the usefulness of cutoff values of fit indices. *Psychological Methods*, *16*(3), 319–336. https://doi.org/10.1037/a0024917

Heise, D. R., & Bohrnstedt, G. W. (1970). Validity, Invalidity, and Reliability. *Sociological Methodology*, *2*, 104–129. https://doi.org/10.2307/270785

Henninger, C. E., Alevizou, P. J., & Oates, C. J. (2016). What is sustainable fashion? *Journal of Fashion Marketing and Management: An International Journal*, *20*(4), 400–416. https://doi.org/10.1108/JFMM-07-2015-0052

Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science*, *43*(1), 115–135. https://doi.org/10.1007/s11747-014-0403-8

Herbes, C., Beuthner, C., & Ramme, I. (2018). Consumer attitudes towards biobased packaging – A cross-cultural comparative study. *Journal of Cleaner Production*, *194*, 203–218. https://doi.org/10.1016/j.jclepro.2018.05.106

Herédia-Colaço, V., & Coelho do Vale, R. (2018). Seize the Day or Save the World? The Importance of Ethical Claims and Product Nature Congruity. *Journal of Business Ethics*, *152*(3), 783–801. https://doi.org/10.1007/s10551-016-3342-0

Hille, S. L., Geiger, C., Loock, M., & Peloza, J. (2018). Best in Class or Simply the Best? The Impact of Absolute versus Relative Ecolabeling Approaches. *Journal of Public Policy & Marketing*, *37*(1), 5–22. https://doi.org/10.1509/jppm.15.030

Hindsley, P., McEvoy, D. M., & Morgan, O. A. (2020). Consumer Demand for Ethical Products and the Role of Cultural Worldviews: The Case of Direct-Trade Coffee. *Ecological Economics*, *177*, 106776. https://doi.org/10.1016/j.ecolecon.2020.106776

Hlavac, M. (2022). *stargazer: Well-Formatted Regression and Summary Statistics Tables* (Version 5.2.3) [Computer software]. Social Policy Institute. https://cran.r-project.org/web/packages/stargazer/index.html

Homburg, C., Koschate, N., & Hoyer, W. D. (2005). Do Satisfied Customers Really Pay More? A Study of the Relationship between Customer Satisfaction and Willingness to Pay: *Journal of Marketing*. https://doi.org/10.1509/jmkg.69.2.84.60760

Homburg, C., Theel, M., & Hohenberg, S. (2020). Marketing Excellence: Nature, Measurement, and Investor Valuations. *Journal of Marketing*, *84*(4), 1–22. https://doi.org/10.1177/0022242920925517

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Hu, Y., Frank, B., & Lu, Z. (2022). Market success through recycling programs: Strategic options, consumer reactions, and contingency factors. *Journal of Cleaner Production*, 131003. https://doi.org/10.1016/j.jclepro.2022.131003

Hunt, S. D., Sparkman, R. D., & Wilcox, J. B. (1982). The Pretest in Survey Research: Issues and Preliminary Findings. *Journal of Marketing Research*, *19*(2), 269–273. https://doi.org/10.2307/3151627

Hur, E. (2020). Rebirth fashion: Secondhand clothing consumption values and perceived risks. *Journal of Cleaner Production*, *273*, 122951. https://doi.org/10.1016/j.jclepro.2020.122951

Hwang, J. W. (2011). Descartes and the Aristotelian Framework of Sensory Perception1: Descartes and Sensory Perception. *Midwest Studies In Philosophy*, *35*(1), 111–148. https://doi.org/10.1111/j.1475-4975.2011.00226.x

Icenogle, G., & Cauffman, E. (2021). Adolescent decision making: A decade in review. *Journal of Research on Adolescence*, *31*(4), 1006–1022. https://doi.org/10.1111/jora.12608

Insch, G. S., & McBride, J. B. (2004). The impact of country-of-origin cues on consumer perceptions of product quality: A binational test of the decomposed country-of-origin construct. *Journal of Business Research*, *57*(3), 256–265. https://doi.org/10.1016/S0148-2963(02)00323-5

Isen, A. M. (1970). Success, failure, attention, and reaction to others: The warm glow of success. *Journal of Personality and Social Psychology*, *15*(4), 294–301. https://doi.org/10.1037/h0029610

Iweala, S., Spiller, A., & Meyerding, S. (2019). Buy good, feel good? The influence of the warm glow of giving on the evaluation of food items with ethical claims in the U.K. and Germany. *Journal of Cleaner Production*, *215*, 315–328. https://doi.org/10.1016/j.jclepro.2018.12.266

Iyer, E. S., & Reczek, R. W. (2017). The Intersection of Sustainability, Marketing, and Public Policy: Introduction to the Special Section on Sustainability: *Journal of Public Policy & Marketing*, *36*(2). https://doi.org/10.1509/jppm.36.250

Jacobs, K., & Hörisch, J. (2021). The importance of product lifetime labelling for purchase decisions: Strategic implications for corporate sustainability based on a conjoint analysis in Germany. *Business Strategy and the Environment*, *31*(4), 1–17. https://doi.org/10.1002/bse.2954

Jacoby, J., & Kaplan, L. B. (1972). The components of perceived risk. In M. Venkatesa (Ed.), *Proceedings of the Association for Consumer Research.* (pp. 382–393).

Jäger, A.-K., & Weber, A. (2020). Can you believe it? The effects of benefit type versus construal level on advertisement credibility and purchase intention for organic food. *Journal of Cleaner Production*, *257*, 120543. https://doi.org/10.1016/j.jclepro.2020.120543

Janßen, D., & Langen, N. (2017). The bunch of sustainability labels – Do consumers differentiate? *Journal of Cleaner Production*, *143*, 1233–1245. https://doi.org/10.1016/j.jclepro.2016.11.171

Jensen, O. (2004). *Key-Account-Management*. Deutscher Universitätsverlag. https://doi.org/10.1007/978-3-663-10468-1

Jensen, P. B., Laursen, L. N., & Haase, L. M. (2021). Barriers to product longevity: A review of business, product development and user perspectives. *Journal of Cleaner Production*, *313*, 127951. https://doi.org/10.1016/j.jclepro.2021.127951

Jeon, J., Shin, H., Jeong, E. (Lena), & Zhang, X. (2024). The power of endorsement in upcycled food promotion: Investigating the moderating effect of perceived food familiarity. *Journal of Retailing and Consumer Services*, *80*, 103901. https://doi.org/10.1016/j.jretconser.2024.103901

Jobst, L. J., Auerswald, M., & Moshagen, M. (2022). The effect of latent and error non-normality on corrections to the test statistic in structural equation modeling. *Behavior Research Methods*, *54*(5), 2351–2363. https://doi.org/10.3758/s13428-021-01729-9

Johnson, M. D. (1989). On the Nature of Product Attributes and Attribute Relationships. *Advances in Consumer Research*, *16*(1), 598–604.

Johnstone, L., & Lindh, C. (2018). The sustainability-age dilemma: A theory of (un)planned behaviour via influencers. *Journal of Consumer Behaviour*, *17*(1), e127–e139. https://doi.org/10.1002/cb.1693

Jones, A. D., Hoey, L., Blesh, J., Miller, L., Green, A., & Shapiro, L. F. (2016). A Systematic Review of the Measurement of Sustainable Diets. *Advances in Nutrition*, *7*(4), 641–664. https://doi.org/10.3945/an.115.011015

Jorge, E., Lopez-Valeiras, E., & Gonzalez-Sanchez, M. B. (2020). The Importance Given to Food Naturalness Attributes by Millennial University Students. *Sustainability*, *12*(2), Article 2. https://doi.org/10.3390/su12020728

Jorgensen, T. D., Pornprasertmanit, S., Schoemann, A. M., Rosseel, Y., Miller, P., Quick, C., Garnier-Villarreal, M., Selig, J., Boulton, A., Preacher, K., Coffman, D., Rhemtulla, M., Robitzsch, A., Enders, C., Arslan, R., Clinton, B., Panko, P., Merkle, E., Chesnut, S., … Johnson, A. R. (2022). *semTools: Useful Tools for Structural Equation Modeling* (Version 0.5-6) [Computer software]. https://cran.r-project.org/web/packages/semTools/index.html

Kaczorowska, J., Prandota, A., Rejman, K., Halicka, E., & Tul-Krzyszczuk, A. (2021). Certification Labels in Shaping Perception of Food Quality—Insights from Polish and Belgian Urban Consumers. *Sustainability*, *13*(2), Article 2. https://doi.org/10.3390/su13020702

Kahraman, A., & Kazançoğlu, İ. (2019). Understanding consumers' purchase intentions toward natural-claimed products: A qualitative research in personal care products. *Business Strategy and the Environment*, *28*(6), 1218–1233. https://doi.org/10.1002/bse.2312

Kaiser, H. F. (1970). A second generation little jiffy. *Psychometrika*, *35*(4), 401–415. https://doi.org/10.1007/BF02291817

Kaiser, H. F., & Rice, J. (1974). Little Jiffy, Mark Iv. *Educational and Psychological Measurement*, *34*(1), 111–117. https://doi.org/10.1177/001316447403400115

Kam, C. C. S. (2023). Why Do Regular and Reversed Items Load on Separate Factors? Response Difficulty vs. Item Extremity. *Educational and Psychological Measurement*, *83*(6), 1085–1112. https://doi.org/10.1177/00131644221143972

Kamalul Ariffin, S., Mohan, T., & Goh, Y.-N. (2018). Influence of consumers' perceived risk on consumers' online purchase intention. *Journal of Research in Interactive Marketing*, *12*(3), 309–327. https://doi.org/10.1108/JRIM-11-2017-0100

Karana, E. (2012). Characterization of 'natural' and 'high-quality' materials to improve perception of bio-plastics. *Journal of Cleaner Production*, *37*, 316–325. https://doi.org/10.1016/j.jclepro.2012.07.034

Kassambara, A. (2023). *rstatix: Pipe-Friendly Framework for Basic Statistical Tests* (Version R package version 0.7.2) [Computer software]. https://CRAN.R-project.org/package=rstatix

Kassie, B. A., & Rhee, H. (2023). The Effect of Brand Familiarity on Green Claim Skepticism in Distribution Channel. *Journal of Distribution Science*, *21*(6), 51–68. https://doi.org/10.15722/JDS.21.06.202306.51

Kasulaitis, B., Babbitt, C. W., & Tyler, A. C. (2021). The role of consumer preferences in reducing material intensity of electronic products. *Journal of Industrial Ecology*, *25*(2), 435–447. https://doi.org/10.1111/jiec.13052

Kelley, K. (2023). *MBESS: The MBESS R Package* (Version R package version 4.9.3) [Computer software]. https://CRAN.R-project.org/package=MBESS

Ketelsen, M., Janssen, M., & Hamm, U. (2020). Consumers' response to environmentally-friendly food packaging—A systematic review. *Journal of Cleaner Production*, *254*, 120123. https://doi.org/10.1016/j.jclepro.2020.120123

Kirmani, A., & Rao, A. R. (2000). No Pain, No Gain: A Critical Review of the Literature on Signaling Unobservable Product Quality. *Journal of Marketing*, *64*(2), 66–79. https://doi.org/10.1509/jmkg.64.2.66.18000

Kmetty, Z., & Stefkovics, Á. (2021). Assessing the effect of questionnaire design on unit and item-nonresponse: Evidence from an online experiment. *International Journal of Social Research Methodology*, *25*(2), 659–672. https://doi.org/10.1080/13645579.2021.1929714

Knoth, K., Fufa, S. M., & Seilskjær, E. (2022). Barriers, success factors, and perspectives for the reuse of construction products in Norway. *Journal of Cleaner Production*, *337*, 130494. https://doi.org/10.1016/j.jclepro.2022.130494

Koller, M., Floh, A., & Zauner, A. (2011). Further insights into perceived value and consumer loyalty: A "Green" perspective. *Psychology & Marketing*, *28*(12), 1154–1176. https://doi.org/10.1002/mar.20432

Komsta, L., & Novomestky, F. (2022). *moments: Moments, Cumulants, Skewness, Kurtosis and Related Tests* (Version R package version 0.14.1) [Computer software]. https://CRAN.R-project.org/package=moments

Konuk, F. A. (2018). The role of store image, perceived quality, trust and perceived value in predicting consumers' purchase intentions towards organic private label food. *Journal of Retailing and Consumer Services*, *43*, 304–310. https://doi.org/10.1016/j.jretconser.2018.04.011

Korkmaz, S., Goksuluk, D., & Zararsiz, G. (2014). MVN: An R Package for Assessing Multivariate Normality. *The R Journal*, *6*(2), 151–162. https://doi.org/10.32614/RJ-2014-031

Kott, K. (2024, November 6). *Einkommen und Einkommensverteilung*. bpb.de. https://www.bpb.de/kurz-knapp/zahlen-und-fakten/sozialbericht-2024/553205/einkommen-und-einkommensverteilung/

Kreuter, F., Presser, S., & Tourangeau, R. (2008). Social Desirability Bias in CATI, IVR, and Web Surveys: The Effects of Mode and Question Sensitivity. *Public Opinion Quarterly*, *72*(5), 847–865. https://doi.org/10.1093/poq/nfn063

Krosnick, J. A., & Presser, S. (2009). Question and Questionnaire Design. In *Handbook of Survey Research* (2nd ed., p. 81). CA: Elsevier.

Kuah, A. T. H., & Wang, P. (2020). Circular economy and consumer acceptance: An exploratory study in East and Southeast Asia. *Journal of Cleaner Production*, *247*, 119097. https://doi.org/10.1016/j.jclepro.2019.119097

Kumar, P., Polonsky, M., Dwivedi, Y. K., & Kar, A. (2021). Green information quality and green brand evaluation: The moderating effects of eco-label credibility and consumer

knowledge. *European Journal of Marketing*, *55*(7), 2037–2071. https://doi.org/10.1108/EJM-10-2019-0808

Kumar, P., & Polonsky, M. J. (2019). In-store experience quality and perceived credibility: A green retailer context. *Journal of Retailing and Consumer Services*, *49*, 23–34. https://doi.org/10.1016/j.jretconser.2019.02.022

Kumar, V., & Reinartz, W. (2016). Creating Enduring Customer Value. *Journal of Marketing*, *80*(6), 36–68. https://doi.org/10.1509/jm.15.0414

Kupfer, A.-K., Pähler vor der Holte, N., Kübler, R. V., & Hennig-Thurau, T. (2018). The Role of the Partner Brand's Social Media Power in Brand Alliances. *Journal of Marketing*, *82*(3), 25–44. https://doi.org/10.1509/jm.15.0536

Langan, R., & Kumar, A. (2019). Time versus money: The role of perceived effort in consumers' evaluation of corporate giving. *Journal of Business Research*, *99*, 295–305. https://doi.org/10.1016/j.jbusres.2019.02.016

Larceneux, F., Benoit-Moreau, F., & Renaudin, V. (2012). Why Might Organic Labels Fail to Influence Consumer Choices? Marginal Labelling and Brand Equity Effects. *Journal of Consumer Policy*, *35*(1), 85–104. https://doi.org/10.1007/s10603-011-9186-1

Larson, P. D. (1994). Buyer-Supplier Co-operation, Product Quality and Total Costs. *International Journal of Physical Distribution & Logistics Management*, *24*(6), 4–10. https://doi.org/10.1108/09600039410066123

Laufer, W. S. (2003). Social Accountability and Corporate Greenwashing. *Journal of Business Ethics*, *43*(3), 253–261. https://doi.org/10.1023/A:1022962719299

Lee, E. J., Bae, J., & Kim, K. H. (2020). The effect of environmental cues on the purchase intention of sustainable products. *Journal of Business Research*, *120*, 425–433. https://doi.org/10.1016/j.jbusres.2019.10.048

Lee, K., Ha, S., & Yoon, B. (2023). Effects of corporate social responsibility (CSR) fit and restaurant size: The mediating role of perceived effort. *Journal of Foodservice Business Research*, *26*(3), 478–495. https://doi.org/10.1080/15378020.2022.2028529

Lee, Y., & Colarelli O'Connor, G. (2003). The Impact of Communication Strategy on Launching New Products: The Moderating Role of Product Innovativeness. *Journal of Product Innovation Management*, *20*(1), 4–21. https://doi.org/10.1111/1540-5885.t01-1-201002

Leiner, D. J. (2019). Too Fast, too Straight, too Weird: Non-Reactive Indicators for Meaningless Data in Internet Surveys. *Survey Research Methods*, *13*(3), Article 3. https://doi.org/10.18148/srm/2019.v13i3.7403

Life Cycle Initiative. (2020). *What is Life Cycle Thinking?* https://www.lifecycleinitiative.org/starting-life-cycle-thinking/what-is-life-cycle-thinking/

Lim, K. H., & Reed, M. (2020). Do ecolabels cheapen wines? *Journal of Cleaner Production*, *245*, 118696. https://doi.org/10.1016/j.jclepro.2019.118696

Lin, Y.-C., & Chang, C. A. (2012). Double Standard: The Role of Environmental Consciousness in Green Product Usage. *Journal of Marketing*, *76*(5), 125–134. https://doi.org/10.2307/41714513

Lindeman, M., & Väänänen, M. (2000). Measurement of ethical food choice motives. *Appetite*, *34*(1), 55–59. https://doi.org/10.1006/appe.1999.0293

Liobikienė, G., & Bernatonienė, J. (2017). Why determinants of green purchase cannot be treated equally? The case of green cosmetics: Literature review. *Journal of Cleaner Production*, *162*, 109–120. https://doi.org/10.1016/j.jclepro.2017.05.204

Lombart, C., Millan, E., Normand, J.-M., Verhulst, A., Labbé-Pinlon, B., & Moreau, G. (2020). Effects of physical, non-immersive virtual, and immersive virtual store environments on consumers' perceptions and purchase behavior. *Computers in Human Behavior*, *110*, 106374. https://doi.org/10.1016/j.chb.2020.106374

Long, J. S., & Ervin, L. H. (2000). Using Heteroscedasticity Consistent Standard Errors in the Linear Regression Model. *The American Statistician*, *54*(3), 217–224. https://doi.org/10.2307/2685594

Lopes, J. M., Gomes, S., & Trancoso, T. (2024). Navigating the green maze: Insights for businesses on consumer decision-making and the mediating role of their environmental concerns. *Sustainability Accounting, Management and Policy Journal*, *15*(4), 861–883. https://doi.org/10.1108/SAMPJ-07-2023-0492

Lopes, J. M., Pinho, M., & Gomes, S. (2024). From green hype to green habits: Understanding the factors that influence young consumers' green purchasing decisions. *Business Strategy and the Environment*, *33*(3), 2432–2444. https://doi.org/10.1002/bse.3602

Lou, X., Chi, T., Janke, J., & Desch, G. (2022). How Do Perceived Value and Risk Affect Purchase Intention toward Second-Hand Luxury Goods? An Empirical Study of U.S. Consumers. *Sustainability*, *14*(18), Article 18. https://doi.org/10.3390/su141811730

Lu, I. R. R., & Kwan, E. (2023). An investigation of two remedial measures for retailers to address the impact of disease threat on sustainable consumption: A moderated moderated mediation model. *Journal of Retailing and Consumer Services*, *72*, 103254. https://doi.org/10.1016/j.jretconser.2023.103254

Luchs, M. G., Naylor, R. W., Irwin, J. R., & Raghunathan, R. (2010). The Sustainability Liability: Potential Negative Effects of Ethicality on Product Preference. *Journal of Marketing*, *74*(5), 18–31. https://doi.org/10.1509/jmkg.74.5.018

Lüdecke, D., Ben-Shachar, M. S., Patil, I., Waggoner, P., & Makowski, D. (2021). performance: An R Package for Assessment, Comparison and Testing of Statistical Models. *Journal of Open Source Software*, *6*(60), 3139. https://doi.org/10.21105/joss.03139

Lv, J., Liu, X., & Cheng, S. (2021). The Impact of Remanufactured Products' Similarity on Purchase Intention of New Products. *Sustainability*, *13*(4), Article 4. https://doi.org/10.3390/su13041825

Lyon, T. P., & Montgomery, A. W. (2015). The Means and End of Greenwash. *Organization & Environment*, *28*(2), 223–249. https://doi.org/10.1177/1086026615575332

Maccioni, L., Borgianni, Y., & Basso, D. (2019). Value Perception of Green Products: An Exploratory Study Combining Conscious Answers and Unconscious Behavioral Aspects. *Sustainability*, *11*(5), Article 5. https://doi.org/10.3390/su11051226

Macdonald, E. K., Kleinaltenkamp, M., & Wilson, H. N. (2016). How Business Customers Judge Solutions: Solution Quality and Value in Use. *Journal of Marketing*, *80*(3), 96–120. https://doi.org/10.1509/jm.15.0109

Magnier, L., & Crié, D. (2015). Communicating packaging eco-friendliness: An exploration of consumers' perceptions of eco-designed packaging. *International Journal of Retail & Distribution Management*, *43*(4/5), 350–366. https://doi.org/10.1108/IJRDM-04-2014-0048

Magnier, L., Mugge, R., & Schoormans, J. (2019). Turning ocean garbage into products – Consumers' evaluations of products made of recycled ocean plastic. *Journal of Cleaner Production*, *215*, 84–98. https://doi.org/10.1016/j.jclepro.2018.12.246 0959-6526

Magnier, L., Schoormans, J., & Mugge, R. (2016). Judging a product by its cover: Packaging sustainability and perceptions of quality in food products. *Food Quality and Preference*, *53*, 132–142. https://doi.org/10.1016/j.foodqual.2016.06.006

Mai, R., Hoffmann, S., Lasarov, W., & Buhs, A. (2019). Ethical Products = Less Strong: How Explicit and Implicit Reliance on the Lay Theory Affects Consumption Behaviors. *Journal of Business Ethics*, *158*(3), 659–677. https://doi.org/10.1007/s10551-017-3669-1

Makov, T., & Fitzpatrick, C. (2021). Is repairability enough? Big data insights into smartphone obsolescence and consumer interest in repair. *Journal of Cleaner Production*, *313*, 127561. https://doi.org/10.1016/j.jclepro.2021.127561

Malik, M. A. A., Mustapha, M. F., Sobri, N. M., Razak, N. F. A., Zaidi, M. N. M., Shukri, A. A., & Sham, M. A. L. Z. (2021). Optimal Reliability and Validity of Measurement Model in Confirmatory Factor Analysis: Different Likert Point Scale Experiment. *Journal of Contemporary Issues and Thought*, *11*, 105–112. https://doi.org/10.37134/jcit.vol11.9.2021

Mancini, M. C., Menozzi, D., Donati, M., Biasini, B., Veneziani, M., & Arfini, F. (2019). Producers' and Consumers' Perception of the Sustainability of Short Food Supply Chains: The Case of Parmigiano Reggiano PDO. *Sustainability*, *11*(3), Article 3. https://doi.org/10.3390/su11030721

Mangiafico, S. S. (2024). *rcompanion: Functions to Support Extension Education Program Evaluation.* (Version R package version 2.4.36) [Computer software]. Rutgers Cooperative Extension. https://CRAN.R-project.org/package=rcompanion

Margariti, K. (2021). "White" Space and Organic Claims on Food Packaging: Communicating Sustainability Values and Affecting Young Adults' Attitudes and Purchase Intentions. *Sustainability*, *13*(19), Article 19. https://doi.org/10.3390/su131911101

Marsh, H. W., Balla, J. R., & McDonald, R. P. (1988). Goodness-of-fit indexes in confirmatory factor analysis: The effect of sample size. *Psychological Bulletin*, *103*(3), 391–410. https://doi.org/10.1037/0033-2909.103.3.391

Martínez, M. P., Cremasco, C. P., Gabriel Filho, L. R. A., Braga Junior, S. S., Bednaski, A. V., Quevedo-Silva, F., Correa, C. M., da Silva, D., & Moura-Leite Padgett, R. C. (2020). Fuzzy inference system to study the behavior of the green consumer facing the perception of greenwashing. *Journal of Cleaner Production*, *242*, 116064. https://doi.org/10.1016/j.jclepro.2019.03.060

Marzouk, O. A., & Mahrous, A. A. (2020). Sustainable Consumption Behavior of Energy and Water-Efficient Products in a Resource-Constrained Environment. *Journal of Global Marketing*, *33*(5), 335–353. https://doi.org/10.1080/08911762.2019.1709005

Matsumoto, M., Chinen, K., & Endo, H. (2017). Comparison of U.S. and Japanese Consumers' Perceptions of Remanufactured Auto Parts. *Journal of Industrial Ecology*, *21*(4), 966–979. https://doi.org/10.1111/jiec.12478

Matsumoto, M., Chinen, K., & Endo, H. (2018). Paving the way for sustainable remanufacturing in Southeast Asia: An analysis of auto parts markets. *Journal of Cleaner Production*, *205*, 1029–1041. https://doi.org/10.1016/j.jclepro.2018.09.074

Mattila, A. S., & Wirtz, J. (2001). Congruency of scent and music as a driver of in-store evaluations and behavior. *Journal of Retailing*, *77*(2), 273. https://doi.org/10.1016/S0022-4359(01)00042-2

McCrum-Gardner, E. (2008). Which is the correct statistical test to use? *British Journal of Oral and Maxillofacial Surgery*, *46*(1), 38–41. https://doi.org/10.1016/j.bjoms.2007.09.002

McDaniel, C., & Baker, R. C. (1977). Convenience Food Packaging and the Perception of Product Quality. *Journal of Marketing*, *41*(4), 57–58. https://doi.org/10.2307/1250234

McDonald, R. P. (1999). *Test Theory: A Unified Treatment*. Psychology Press. https://doi.org/10.4324/9781410601087

McDonald, S. M. (2012). Perception: A Concept Analysis. *International Journal of Nursing Knowledge*, *23*(1), 2–9. https://doi.org/10.1111/j.2047-3095.2011.01198.x

McIntyre, F., & Castleberry, S. (1992). Consumers' Quality Evaluation Process. *Journal of Applied Business Research*, *8*, 74–82. https://doi.org/10.19030/jabr.v8i3.6147

Meyer, D., Zeileis, A., Hornik, K., & Friendly, M. (2024). *vcd: Visualizing Categorical Data* (Version R package version 1.4-13) [Computer software]. https://cran.r-project.org/web/packages/vcd/index.html

Michaud, C., & Llerena, D. (2011). Green consumer behaviour: An experimental analysis of willingness to pay for remanufactured products. *Business Strategy and the Environment*, *20*(6), 408–420. https://doi.org/10.1002/bse.703

Michel, F., & Siegrist, M. (2019). How should importance of naturalness be measured? A comparison of different scales. *Appetite*, *140*, 298–304. https://doi.org/10.1016/j.appet.2019.05.019

Michon, R., Chebat, J.-C., & Turley, L. W. (2005). Mall atmospherics: The interaction effects of the mall environment on shopping behavior. *Journal of Business Research*, *58*(5), 576–583. https://doi.org/10.1016/j.jbusres.2003.07.004

Miklós-Thal, J., & Zhang, J. (2013). (De)marketing to Manage Consumer Quality Inferences. *Journal of Marketing Research*, *50*(1), 55–69. https://doi.org/10.1509/jmr.11.0432

Molina-Castillo, F.-J., Calantone, R. J., Stanko, M. A., & Munuera-Aleman, J.-L. (2013). Product Quality as a Formative Index: Evaluating an Alternative Measurement Approach. *Journal of Product Innovation Management*, *30*(2), 380–398. https://doi.org/10.1111/j.1540-5885.2012.01005.x

Mollenkopf, D. A., Peinkofer, S. T., & Chu, Y. J. (2022). Supply chain transparency: Consumer reactions to incongruent signals. *Journal of Operations Management*, *68*(4), 306–327. https://doi.org/10.1002/joom.1180

Monnot, E., Parguel, B., & Reniou, F. (2015). Consumer responses to elimination of overpackaging on private label products. *International Journal of Retail & Distribution Management*, *43*(4/5), 329–349. https://doi.org/10.1108/IJRDM-03-2014-0036

Montgomery, D. C., Peck, E. A., & Vining, G. G. (2021). *Introduction to Linear Regression Analysis*. John Wiley & Sons.

Moons, I., Barbarossa, C., & De Pelsmacker, P. (2018). The Determinants of the Adoption Intention of Eco-friendly Functional Food in Different Market Segments. *Ecological Economics*, *151*, 151–161. https://doi.org/10.1016/j.ecolecon.2018.05.012

Morales, A. C. (2005). Giving Firms an "E" for Effort: Consumer Responses to High-Effort Firms. *Journal of Consumer Research*, *31*(4), 806–812. https://doi.org/10.1086/426615

Mugge, R., Jockin, B., & Bocken, N. (2017). How to sell refurbished smartphones? An investigation of different customer groups and appropriate incentives. *Journal of Cleaner Production*, *147*, 284–296. https://doi.org/10.1016/j.jclepro.2017.01.111

Mukendi, A., Davies, I., Glozer, S., & McDonagh, P. (2020). Sustainable fashion: Current and future research directions. *European Journal of Marketing*, *54*(11), 2873–2909. https://doi.org/10.1108/EJM-02-2019-0132

Munten, P., & Vanhamme, J. (2023). To reduce waste, have it repaired! The quality signaling effect of product repairability. *Journal of Business Research*, *156*, 113457. https://doi.org/10.1016/j.jbusres.2022.113457

Myers, D. G., & DeWall, C. N. (2015). *Psychology* (Eleventh edition). Worth Publishers/A MacMillian Education Imprint.

Nazlı, T. (2021). Repair motivation and barriers model: Investigating user perspectives related to product repair towards a circular economy. *Journal of Cleaner Production*, *289*, 125644. https://doi.org/10.1016/j.jclepro.2020.125644

Nekmahmud, M., & Fekete-Farkas, M. (2020). Why Not Green Marketing? Determinates of Consumers' Intention to Green Purchase Decision in a New Developing Nation. *Sustainability*, *12*(19), Article 19. https://doi.org/10.3390/su12197880

Nenonen, S., Storbacka, K., & Windahl, C. (2019). Capabilities for market-shaping: Triggering and facilitating increased value creation. *Journal of the Academy of Marketing Science*, *47*(4), 617–639. https://doi.org/10.1007/s11747-019-00643-z

Newell, S. J., & Goldsmith, R. E. (2001). The development of a scale to measure perceived corporate credibility. *Journal of Business Research*, *52*(3), 235–247. https://doi.org/10.1016/S0148-2963(99)00104-6

Newman, G. E., Gorlin, M., & Dhar, R. (2014). When Going Green Backfires: How Firm Intentions Shape the Evaluation of Socially Beneficial Product Enhancements. *Journal of Consumer Research*, *41*(3), 823–839. https://doi.org/10.1086/677841

Ng, P. F., Butt, M. M., Khong, K. W., & Ong, F. S. (2014). Antecedents of Green Brand Equity: An Integrated Approach. *Journal of Business Ethics*, *121*(2), 203–215. https://doi.org/10.1007/s10551-013-1689-z

Ng, P. M. L., Cheung, C. T. Y., Lit, K. K., Wan, C., & Choy, E. T. K. (2024). Green consumption and sustainable development: The effects of perceived values and

motivation types on green purchase intention. *Business Strategy and the Environment*, *33*(2), 1024–1039. https://doi.org/10.1002/bse.3535

Nguyen, A. T., Parker, L., Brennan, L., & Lockrey, S. (2020). A consumer definition of eco-friendly packaging. *Journal of Cleaner Production*, *252*, 119792. https://doi.org/10.1016/j.jclepro.2019.119792

Nguyen, T. T. H., Yang, Z., Nguyen, N., Johnson, L. W., & Cao, T. K. (2019). Greenwash and Green Purchase Intention: The Mediating Role of Green Skepticism. *Sustainability*, *11*(9), 2653. https://doi.org/10.3390/su11092653

Nichols, B. S., Kirchoff, J., Stolze, H., & Confente, I. (2023). When brands behave badly: Signaling and spillover effects of unethical behavior in the context of triple bottom line sustainability. *Journal of Product & Brand Management*, *32*(6), 908–926. https://doi.org/10.1108/JPBM-07-2021-3569]

Nichols, B. S., Stolze, H., & Kirchoff, J. F. (2019). Spillover effects of supply chain news on consumers' perceptions of product quality: An examination within the triple bottom line. *Journal of Operations Management*, *65*(6), 536–559. https://doi.org/10.1002/joom.1033

Nielsen, P. H., & Wenzel, H. (2002). Integration of environmental aspects in product development: A stepwise procedure based on quantitative life cycle assessment. *Journal of Cleaner Production*, *10*(3), 247–257. https://doi.org/10.1016/S0959-6526(01)00038-5

Norman, J. (2002). Two visual systems and two theories of perception: An attempt to reconcile the constructivist and ecological approaches. *Behavioral and Brain Sciences*, *25*(1), 73–96. https://doi.org/10.1017/S0140525X0200002X

Normann, A., Röding, M., & Wendin, K. (2019). Sustainable Fruit Consumption: The Influence of Color, Shape and Damage on Consumer Sensory Perception and Liking of Different Apples. *Sustainability*, *11*(17), Article 17. https://doi.org/10.3390/su11174626

Nunes, J. C., Ordanini, A., & Giambastiani, G. (2021). The Concept of Authenticity: What It Means to Consumers. *Journal of Marketing*, *85*(4), 1–20. https://doi.org/10.1177/0022242921997081

Nunes, P. A. L. D., & Schokkaert, E. (2003). Identifying the warm glow effect in contingent valuation. *Journal of Environmental Economics and Management*, *45*(2), 231–245. https://doi.org/10.1016/S0095-0696(02)00051-7

O'brien, R. M. (2007). A Caution Regarding Rules of Thumb for Variance Inflation Factors. *Quality & Quantity*, *41*(5), 673–690. https://doi.org/10.1007/s11135-006-9018-6

OpenAI. (2024). *ChatGPT* [Large language model]. OpenAI. https://chat.openai.com/chat

Oppong, F. B., & Agbedra, S. Y. (2016). Assessing Univariate and Multivariate Normality, A Guide For Non-Statisticians. *Mathematical Theory and Modelling*, *6*(2), 26–33.

Osburg, V.-S., Yoganathan, V., Bartsch, F., Diallo, M. F., & Liu, H. (2024). How Sustainable Luxury Influences Product Value Perceptions and Behavioral Intentions: A Comparative Study of Emerging vs. Developed Markets. *Journal of Business Ethics*, *191*(4), 713–738. https://doi.org/10.1007/s10551-024-05661-8

Osburg, V.-S., Yoganathan, V., Brueckner, S., & Toporowski, W. (2019). How detailed product information strengthens eco-friendly consumption. *Management Decision*, *58*(6), 1084–1099. https://doi.org/10.1108/MD-10-2017-1012

Otto, S., Strenger, M., Maier-Nöth, A., & Schmid, M. (2021). Food packaging and sustainability – Consumer perception vs. correlated scientific facts: A review. *Journal of Cleaner Production*, *298*, 126733. https://doi.org/10.1016/j.jclepro.2021.126733

Ovchinnikov, A. (2011). Revenue and Cost Management for Remanufactured Products. *Production and Operations Management*, *20*(6), 824–840. https://doi.org/10.1111/j.1937-5956.2010.01214.x

Owen, R., Brennan, G., & Lyon, F. (2018). Enabling investment for the transition to a low carbon economy: Government policy to finance early stage green innovation. *Current Opinion in Environmental Sustainability*, *31*, 137–145. https://doi.org/10.1016/j.cosust.2018.03.004

Ozanne, L. K., Phipps, M., Weaver, T., Carrington, M., Luchs, M., Catlin, J., Gupta, S., Santos, N., Scott, K., & Williams, J. (2016). Managing the Tensions at the Intersection of the Triple Bottom Line: A Paradox Theory Approach to Sustainability Management. *Journal of Public Policy & Marketing*, *35*(2), 249–261. https://doi.org/10.1509/jppm.15.143

Pahlevi, M. R., & Suhartanto, D. (2020). The integrated model of green loyalty: Evidence from eco-friendly plastic products. *Journal of Cleaner Production*, *257*, 120844. https://doi.org/10.1016/j.jclepro.2020.120844

Palmatier, R. W., Houston, M. B., & Hulland, J. (2018). Review articles: Purpose, process, and structure. *Journal of the Academy of Marketing Science*, *46*(1), 1–5. https://doi.org/10.1007/s11747-017-0563-4

Pancer, E., McShane, L., & Noseworthy, T. (2017). Isolated Environmental Cues and Product Efficacy Penalties: The Color Green and Eco-labels. *Journal of Business Ethics*, *143*(1), 159–177. https://doi.org/10.1007/s10551-015-2764-4

Paparoidamis, N. G., & Tran, H. T. T. (2019). Making the world a better place by making better products: Eco-friendly consumer innovativeness and the adoption of eco-innovations. *European Journal of Marketing*, *53*(8), 1546–1584. https://doi.org/10.1108/EJM-11-2017-0888

Paparoidamis, N. G., Tran, T. T. H., Leonidou, L. C., & Zeriti, A. (2019). Being Innovative While Being Green: An Experimental Inquiry into How Consumers Respond to Eco-Innovative Product Designs. *Journal of Product Innovation Management*, *36*(6), 824–847. https://doi.org/10.1111/jpim.12509

Pappu, R., & Quester, P. G. (2016). How does brand innovativeness affect brand loyalty? *European Journal of Marketing*, *50*(1/2), 2–28. https://doi.org/10.1108/EJM-01-2014-0020

Park, H., Lalwani, A. K., & Silvera, D. H. (2020). The Impact of Resource Scarcity on Price-Quality Judgments. *Journal of Consumer Research*, *46*(6), 1110–1124. https://doi.org/10.1093/jcr/ucz031

Park, T., & Kim, J. (2024). Too innovative to be recycled: The role of perceived innovativeness in recycled product advertisements. *Corporate Social Responsibility and Environmental Management*, *31*(6), 5850–5862. https://doi.org/10.1002/csr.2897

Patrício, M., Ferreira, F., Oliveiros, B., & Caramelo, F. (2017). Comparing the performance of normality tests with ROC analysis and confidence intervals. *Communications in Statistics - Simulation and Computation*, *46*(10), 7535–7551. https://doi.org/10.1080/03610918.2016.1241410

Peattie, K. (1999). Trappings versus substance in the greening of marketing planning. *Journal of Strategic Marketing*, *7*(2), 131–148. https://doi.org/10.1080/096525499346486

Pedersen, E. R., & Neergaard, P. (2006). Caveat emptor – let the buyer beware! Environmental labelling and the limitations of 'green' consumerism. *Business Strategy and the Environment*, *15*(1), 15–29. https://doi.org/10.1002/bse.434

Perez, D., Stockheim, I., Tevet, D., & Matan Rubin, M. (2020). Consumers value manufacturer sincerity: The effect of central eco-friendly attributes on luxury product evaluations. *Journal of Cleaner Production*, *267*, 122132. https://doi.org/10.1016/j.jclepro.2020.122132

Peter, J. P. (1981). Construct Validity: A Review of Basic Issues and Marketing Practices. *Journal of Marketing Research*, *18*(2), 133–145. https://doi.org/10.2307/3150948

Petersen, M., & Brockhaus, S. (2017). Dancing in the dark: Challenges for product developers to improve and communicate product sustainability. *Journal of Cleaner Production*, *161*, 345–354. https://doi.org/10.1016/j.jclepro.2017.05.127

Peytchev, A., Couper, M. P., McCabe, S. E., & Crawford, S. D. (2006). Web Survey Design. *Public Opinion Quarterly*, *70*(4), 596–607. https://doi.org/10.1093/poq/nfl028

Polites, G. L., Roberts, N., & Thatcher, J. (2012). Conceptualizing models using multidimensional constructs: A review and guidelines for their use. *European Journal of Information Systems*, *21*(1), 22–48. https://doi.org/10.1057/ejis.2011.10

Poushter, J., & Huang, C. (2020). *Despite Pandemic, Many Europeans Still See Climate Change as Greatest Threat to Their Countries*. Pew Research Center. https://www.pewresearch.org/global/2020/09/09/despite-pandemic-many-europeans-still-see-climate-change-as-greatest-threat-to-their-countries/

Prados-Peña, M. B., Gálvez-Sánchez, F. J., & García-López, A. (2023). Moving toward sustainable development: Social, economic and environmental value as antecedents of purchase intention in the sustainable crafts sector. *Sustainable Development*, *31*(4), 3024–3037. https://doi.org/10.1002/sd.2567

Prata, J. C., da Costa, J. P., Lopes, I., Duarte, A. C., & Rocha-Santos, T. (2020). Environmental exposure to microplastics: An overview on possible human health effects. *Science of The Total Environment*, *702*, 134455. https://doi.org/10.1016/j.scitotenv.2019.134455

Preziosi, M., Tourais, P., Acampora, A., Videira, N., & Merli, R. (2019). The role of environmental practices and communication on guest loyalty: Examining EU-Ecolabel in Portuguese hotels. *Journal of Cleaner Production*, *237*, 117659. https://doi.org/10.1016/j.jclepro.2019.117659

Prolific. (2024). www.prolific.com

Pula, K., Parks, C. D., & Ross, C. F. (2014). Regulatory focus and food choice motives. Prevention orientation associated with mood, convenience, and familiarity. *Appetite*, *78*, 15–22. https://doi.org/10.1016/j.appet.2014.02.015

R Core Team. (2024). *R: A Language and Environment for Statistical Computing* [Computer software]. R Foundation for Statistical Computing. https://www.R-project.org/

Rahinel, R., & Nelson, N. M. (2016). When Brand Logos Describe the Environment: Design Instability and the Utility of Safety-Oriented Products. *Journal of Consumer Research*, *43*(3), 478–496. https://doi.org/10.1093/jcr/ucw039

Rahman, S. ur, & Nguyen-Viet, B. (2023). Towards sustainable development: Coupling green marketing strategies and consumer perceptions in addressing greenwashing. *Business Strategy and the Environment*, *32*(4), 2420–2433. https://doi.org/10.1002/bse.3256

Raj, S., Singh, A., & Lascu, D.-N. (2023). Green smartphone purchase intentions: A conceptual framework and empirical investigation of Indian consumers. *Journal of Cleaner Production*, *403*, 136658. https://doi.org/10.1016/j.jclepro.2023.136658

Ramesh, K., Saha, R., Goswami, S., Sekar, & Dahiya, R. (2019). Consumer's response to CSR activities: Mediating role of brand image and brand attitude. *Corporate Social Responsibility and Environmental Management*, *26*(2), 377–387. https://doi.org/10.1002/csr.1689

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

Rao, A. R. (2005). The Quality of Price as a Quality Cue. *Journal of Marketing Research*, *42*(4), 401–405. https://doi.org/10.1509/jmkr.2005.42.4.401

Ray, S., Danks, N. P., & Calero Valdez, A. (2024). *seminr: Building and Estimating Structural Equation Models* (Version R package version 2.3.4) [Computer software]. https://CRAN.R-project.org/package=seminr

Razali, N. M., & Wah, Y. B. (2011). Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests. *Journal of Statistical Modeling and Analytics*, *2*(1), 21–33.

Reeves, C. A., & Bednar, D. A. (1994). Defining Quality: Alternatives and Implications. *The Academy of Management Review*, *19*(3), 419–445. https://doi.org/10.2307/258934

Reinhardt, F. L. (1998). Environmental Product Differentiation: IMPLICATIONS FOR CORPORATE STRATEGY. *California Management Review*, *40*(4), 43–73. https://doi.org/10.2307/41165964

Revelle, W. (2024). *psych: Procedures for Psychological, Psychometric, and Personality Research* (Version R package version 2.4.6) [Computer software]. Northwestern University. https://CRAN.R-project.org/package=psych

Reynolds, W. M. (1982). Development of reliable and valid short forms of the marlowe-crowne social desirability scale. *Journal of Clinical Psychology*, *38*(1), 119–125. https://doi.org/10.1002/1097-4679(198201)38:1<119::AID-JCLP2270380118>3.0.CO;2-I

Ringle, C. M., Sarstedt, M., Sinkovics, N., & Sinkovics, R. R. (2023). A perspective on using partial least squares structural equation modelling in data articles. *Data in Brief*, *48*, 109074. https://doi.org/10.1016/j.dib.2023.109074

Riva, F., Magrizos, S., Rubel, M. R. B., & Rizomyliotis, I. (2022). Green consumerism, green perceived value, and restaurant revisit intention: Millennials' sustainable consumption

with moderating effect of green perceived quality. *Business Strategy and the Environment*, *31*(7), 2807–2819. https://doi.org/10.1002/bse.3048

Rogers, H. A., Deutz, P., & Ramos, T. B. (2021). Repairing the circular economy: Public perception and participant profile of the repair economy in Hull, UK. *Resources, Conservation and Recycling*, *168*, 105447. https://doi.org/10.1016/j.resconrec.2021.105447

Roithner, C., Cencic, O., & Rechberger, H. (2022). Product design and recyclability: How statistical entropy can form a bridge between these concepts - A case study of a smartphone. *Journal of Cleaner Production*, *331*, 129971. https://doi.org/10.1016/j.jclepro.2021.129971

Román, S., Sánchez-Siles, L. M., & Siegrist, M. (2017). The importance of food naturalness for consumers: Results of a systematic review. *Trends in Food Science & Technology*, *67*, 44–57. https://doi.org/10.1016/j.tifs.2017.06.010

Ross, S. M., & Milne, G. R. (2020). Price? Quality? Or Sustainability? Segmenting by Disposition Toward Self-other Tradeoffs Predicts Consumers' Sustainable Decision-Making. *Journal of Business Ethics*, *172*(2), 361–378. https://doi.org/10.1007/s10551-020-04478-5

Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, *48*, 1–36. https://doi.org/10.18637/jss.v048.i02

Rost, Z. (2015). The increasing relevance of product responsibility. *Uwf UmweltWirtschaftsForum*, *23*(4), 299–305. https://doi.org/10.1007/s00550-015-0369-0

Royne, M. B., Martinez, J., Oakley, J., & Fox, A. K. (2012). The effectiveness of benefit type and price endings in green advertising. *Journal of Advertising*, *41*(4), 85–102. https://doi.org/10.1080/00913367.2012.10672459

Rufibach, K. (2009). reporttools: R Functions to Generate LaTeX Tables of Descriptive Statistics. *Journal of Statistical Software*, *31*, 1–7. https://doi.org/10.18637/jss.v031.c01

Ryoo, Y., & Kim, W. (2023). Price-ethicality association: When price discounts inhibit ethical purchasing. *Journal of Business Research*, *169*, 114242. https://doi.org/10.1016/j.jbusres.2023.114242

Saldaña, J. (2013). *The coding manual for qualitative researchers* (2nd ed). SAGE Publications.

Samoggia, A., Perazzolo, C., Kocsis, P., & Del Prete, M. (2019). Community Supported Agriculture Farmers' Perceptions of Management Benefits and Drawbacks. *Sustainability*, *11*(12), Article 12. https://doi.org/10.3390/su11123262

Sánchez-González, I., Gil-Saura, I., & Ruiz-Molina, M.-E. (2021). Does sustainability drive to create store equity? A proposal through image, quality and loyalty. *International Journal of Retail & Distribution Management*, *50*(6), 708–727. https://doi.org/10.1108/IJRDM-10-2020-0389

Sarkar, J. G., Sarkar, A., Dwivedi, Y. K., & Balaji, M. S. (2022). Sweat it for sustainability: Impact of physical activity/exercise on sustainable consumption. *Psychology & Marketing*, *39*(11), 2184–2199. https://doi.org/10.1002/mar.21722

Satorra, A., & Bentler, P. M. (1994). Corrections to test statistics and standard errors in covariance structure analysis. In *Latent variables analysis: Applications for developmental research* (pp. 399–419). SAGE Publications.

Scarpi, D., Russo, I., Confente, I., & Hazen, B. (2021). Individual antecedents to consumer intention to switch to food waste bioplastic products: A configuration analysis. *Industrial Marketing Management*, *93*, 578–590. https://doi.org/10.1016/j.indmarman.2020.09.006

Schacter, D. L., Gilbert, D. T., & Wegner, D. M. (2011). *Psychology* (2nd ed). Worth Publishers.

Schäufele, I., & Hamm, U. (2017). Consumers' perceptions, preferences and willingness-to-pay for wine with sustainability characteristics: A review. *Journal of Cleaner Production*, *147*, 379–394. https://doi.org/10.1016/j.jclepro.2017.01.118

Schindler, R. (2006). The 99 price ending as a signal of a low-price appeal. *Journal of Retailing*, *82*(1), 71–77. https://doi.org/10.1016/j.jretai.2005.11.001

Schwarz, G. (1978). Estimating the Dimension of a Model. *The Annals of Statistics*, *6*(2), 461–464.

Sdrolia, E., & Zarotiadis, G. (2019). A Comprehensive Review for Green Product Term: From Definition to Evaluation. *Journal of Economic Surveys*, *33*(1), 150–178. https://doi.org/10.1111/joes.12268

Sekhon, T. S., & Armstrong Soule, C. A. (2020). Conspicuous anticonsumption: When green demarketing brands restore symbolic benefits to anticonsumers. *Psychology & Marketing*, *37*(2), 278–290. https://doi.org/10.1002/mar.21299

Sen, S., & Bhattacharya, C. B. (2001). Does Doing Good Always Lead to Doing Better? Consumer Reactions to Corporate Social Responsibility. *Journal of Marketing Research*, *38*(2), 225–243. https://doi.org/10.1509/jmkr.38.2.225.18838

Senbabaoglu, E. (2017). The Impact of the Service Innovativeness on Perceived Overall Service Quality, Customer Loyalty and Perceived Customer Value in Shopping Sites. *Marketing and Branding Research*, *4*(4), 371–383. https://doi.org/10.33844/mbr.2017.60438

Sevilla, J., & Townsend, C. (2016). The Space-to-Product Ratio Effect: How Interstitial Space Influences Product Aesthetic Appeal, Store Perceptions, and Product Preference. *Journal of Marketing Research*, *53*(5), 665–681. https://doi.org/10.1509/jmr.13.0601

Shams, R., Alpert, F., & Brown, M. (2015). Consumer perceived brand innovativeness. *European Journal of Marketing*, *49*(9/10), 1589–1615. https://doi.org/10.1108/EJM-05-2013-0240

Shapiro, S. S., Wilk, M. B., & Chen, H. J. (1968). A Comparative Study of Various Tests for Normality. *Journal of the American Statistical Association*, *63*(324), 1343–1372. https://doi.org/10.1080/01621459.1968.10480932

Sharma, A. P. (2021). Consumers' purchase behaviour and green marketing: A synthesis, review and agenda. *International Journal of Consumer Studies*, *45*(6), 1217–1238. https://doi.org/10.1111/ijcs.12722

Sharma, K., Aswal, C., & Paul, J. (2023). Factors affecting green purchase behavior: A systematic literature review. *Business Strategy and the Environment*, *32*(4), 2078–2092. https://doi.org/10.1002/bse.3237

Sheth, J. N., Newman, B. I., & Gross, B. L. (1991). Why we buy what we buy: A theory of consumption values. *Journal of Business Research*, *22*(2), 159–170. https://doi.org/10.1016/0148-2963(91)90050-8

Shi, D., Maydeu-Olivares, A., & DiStefano, C. (2018). The Relationship Between the Standardized Root Mean Square Residual and Model Misspecification in Factor Analysis Models. *Multivariate Behavioral Research*, *53*(5), 676–694. https://doi.org/10.1080/00273171.2018.1476221

Shi, J., Visschers, V. H. M., Bumann, N., & Siegrist, M. (2018). Consumers' climate-impact estimations of different food products. *Journal of Cleaner Production*, *172*, 1646–1653. https://doi.org/10.1016/j.jclepro.2016.11.140

Shmueli, G., Sarstedt, M., Hair, J. F., Cheah, J.-H., Ting, H., Vaithilingam, S., & Ringle, C. M. (2019). Predictive model assessment in PLS-SEM: Guidelines for using PLSpredict. *European Journal of Marketing*, *53*(11), 2322–2347. https://doi.org/10.1108/EJM-02-2019-0189

Siegel, A. F. (2012). Chapter 3 - Histograms: Looking at the Distribution of Data. In A. F. Siegel (Ed.), *Practical Business Statistics (Sixth Edition)* (pp. 35–64). Academic Press. https://doi.org/10.1016/B978-0-12-385208-3.00003-1

Siegrist, M., & Sütterlin, B. (2017). Importance of perceived naturalness for acceptance of food additives and cultured meat. *Appetite*, *113*, 320–326. https://doi.org/10.1016/j.appet.2017.03.019

Singhal, D., Jena, S. K., & Tripathy, S. (2019). Factors influencing the purchase intention of consumers towards remanufactured products: A systematic review and meta-analysis. *International Journal of Production Research*, *57*(23), 7289–7299. https://doi.org/10.1080/00207543.2019.1598590

Skard, S., Jørgensen, S., & Pedersen, L. J. T. (2021). When is Sustainability a Liability, and When Is It an Asset? Quality Inferences for Core and Peripheral Attributes. *Journal of Business Ethics*, *173*(1), 109–132. https://doi.org/10.1007/s10551-019-04415-1

Slakey, T. J. (1961). Aristotle on Sense Perception. *The Philosophical Review*, *70*(4), 470–484. https://doi.org/10.2307/2183608

Smith, M., Love, D. C., Rochman, C. M., & Neff, R. A. (2018). Microplastics in Seafood and the Implications for Human Health. *Current Environmental Health Reports*, *5*(3), 375–386. https://doi.org/10.1007/s40572-018-0206-z

Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research*, *104*, 333–339. https://doi.org/10.1016/j.jbusres.2019.07.039

Sonego, M., Echeveste, M. E. S., & Debarba, H. G. (2022). Repair of electronic products: Consumer practices and institutional initiatives. *Sustainable Production and Consumption*, *30*, 556–565. https://doi.org/10.1016/j.spc.2021.12.031

Sonnenberg, N. C., Erasmus, A. C., & Donoghue, S. (2011). Significance of environmental sustainability issues in consumers' choice of major household appliances in South Africa. *International Journal of Consumer Studies*, *35*(2), 153–163. https://doi.org/10.1111/j.1470-6431.2010.00964.x

Souri, M. E., Sajjadian, F., Sheikh, R., & Sana, S. S. (2018). Grey SERVQUAL method to measure consumers' attitudes towards green products—A case study of Iranian

consumers of LED bulbs. *Journal of Cleaner Production*, *177*, 187–196. https://doi.org/10.1016/j.jclepro.2017.12.105

Staffolani, G., Bentivoglio, D., & Finco, A. (2022). Consumers' Purchasing Determinants Towards Mountain Food Products. *Sustainability*, *14*(14), Article 14. https://doi.org/10.3390/su14148282

Staub, C., Michel, F., Bucher, T., & Siegrist, M. (2020). How do you perceive this wine? Comparing naturalness perceptions of Swiss and Australian consumers. *Food Quality and Preference*, *79*. https://doi.org/10.1016/j.foodqual.2019.103752

Steenis, N. D., van der Lans, I. A., van Herpen, E., & van Trijp, H. C. M. (2018). Effects of sustainable design strategies on consumer preferences for redesigned packaging. *Journal of Cleaner Production*, *205*, 854–865. https://doi.org/10.1016/j.jclepro.2018.09.137

Steenis, N. D., Van Herpen, E., Van Der Lans, I. A., Ligthart, T. N., & Van Trijp, H. C. m. (2017). Consumer response to packaging design: The role of packaging materials and graphics in sustainability perceptions and product evaluations. *Journal of Cleaner Production*, *162*, 286–298. https://doi.org/10.1016/j.jclepro.2017.06.036

Steenkamp, J.-B. E. M. (1990). Conceptual model of the quality perception process. *Journal of Business Research*, *21*(4), 309–333. https://doi.org/10.1016/0148-2963(90)90019-A

Steinemann, M., Schwegler, R., & Spescha, G. (2017). *Grüne Produkte in Deutschland 2017. Marktbeobachtungen für die Umweltpolitik*. Umweltbundesamt. https://www.umweltbundesamt.de/sites/default/files/medien/1410/publikationen/171206_uba_fb_gruneprodukte_bf_low.pdf

Steinhart, Y., Ayalon, O., & Puterman, H. (2013). The effect of an environmental claim on consumers' perceptions about luxury and utilitarian products. *Journal of Cleaner Production*, *53*, 277–286. https://doi.org/10.1016/j.jclepro.2013.04.024

Steptoe, A., Pollard, T. M., & Wardle, J. (1995). Development of a measure of the motives underlying the selection of food: The food choice questionnaire. *Appetite*, *25*(3), 267–284. https://doi.org/10.1006/appe.1995.0061

Stolz, J., & Bautista, R. (2015). Corporate sustainability: Perception and response by older consumers. *International Journal of Consumer Studies*, *39*(4), 343–351. https://doi.org/10.1111/ijcs.12199

Stone, R. N., & Grønhaug, K. (1993). Perceived Risk: Further Considerations for the Marketing Discipline. *European Journal of Marketing*, *27*(3), 39–50. https://doi.org/10.1108/03090569310026637

Strack, F. (1992). "Order Effects" in Survey Research: Activation and Information Functions of Preceding Questions. In N. Schwarz & S. Sudman (Eds.), *Context Effects in Social and Psychological Research* (pp. 23–34). Springer. https://doi.org/10.1007/978-1-4612-2848-6_3

Stremmel, G., Elshiewy, O., & Boztug, Y. (2024). Climate-neutral labeling for climate-friendly vs. climate-harmful food products: Consumer perceptions and implications. *Business Strategy and the Environment*, *33*(5), 4441–4455. https://doi.org/10.1002/bse.3713

Subramaniam, M., & Youndt, M. A. (2005). The Influence of Intellectual Capital on the Types of Innovative Capabilities. *The Academy of Management Journal*, *48*(3), 450–463. https://doi.org/10.5465/AMJ.2005.17407911

Suki, N. M. (2016). Consumer environmental concern and green product purchase in Malaysia: Structural effects of consumption values. *Journal of Cleaner Production*, *132*, 204–214. https://doi.org/10.1016/j.jclepro.2015.09.087

Sun, H., Teh, P., & Linton, J. D. (2018). Impact of environmental knowledge and product quality on student attitude toward products with recycled/remanufactured content: Implications for environmental education and green manufacturing. *Business Strategy & the Environment*, *27*(7), 935–945. https://doi.org/10.1002/bse.2043

Sun, J. J., Bellezza, S., & Paharia, N. (2021). Buy Less, Buy Luxury: Understanding and Overcoming Product Durability Neglect for Sustainable Consumption. *Journal of Marketing*, *85*(3), 28–43. https://doi.org/10.1177/0022242921993172

Sundar, A., & Kellaris, J. (2017). How Logo Colors Influence Shoppers' Judgments of Retailer Ethicality: The Mediating Role of Perceived Eco-Friendliness. *Journal of Business Ethics*, *146*(3), 685–701. https://doi.org/10.1007/s10551-015-2918-4

Suri, R. (1996). *An investigation of the effects of time pressure on consumers' perception of price, quality, and value.* University of Illinois.

SurveyCircle. (2024). *Forschungswebseite SurveyCircle. Veröffentlicht 2016*. https://www.surveycircle.com

SurveySwap B.V. (2024). *SurveySwap*. https://surveyswap.io/de/

Sutton, J. (2000). Symposium on Descartes on perceptual cognition. In *Descartes' natural philosophy* (pp. 524–527). Routledge.

Sweeney, J. C., & Soutar, G. N. (2001). Consumer perceived value: The development of a multiple item scale. *Journal of Retailing*, *77*(2), 203–220. https://doi.org/10.1016/S0022-4359(01)00041-0

Taber, K. S. (2018). The Use of Cronbach's Alpha When Developing and Reporting Research Instruments in Science Education. *Research in Science Education*, *48*(6), 1273–1296. https://doi.org/10.1007/s11165-016-9602-2

Taufik, D. (2018). Prospective "warm-glow" of reducing meat consumption in China: Emotional associations with intentions for meat consumption curtailment and consumption of meat substitutes. *Journal of Environmental Psychology*, *60*, 48–54. https://doi.org/10.1016/j.jenvp.2018.10.004

Taufik, D., Bolderdijk, J. W., & Steg, L. (2016). Going green? The relative importance of feelings over calculation in driving environmental intent in the Netherlands and the United States. *Energy Research & Social Science*, *22*, 52–62. https://doi.org/10.1016/j.erss.2016.08.012

Techawachirakul, M., Pathak, A., Motoki, K., & Calvert, G. A. (2023). Negative halo effects of sustainable packaging. *Psychology & Marketing*, *40*(12), 2627–2641. https://doi.org/10.1002/mar.21891

Testa, F., Iraldo, F., Vaccari, A., & Ferrari, E. (2015). Why Eco-labels can be Effective Marketing Tools: Evidence from a Study on Italian Consumers. *Business Strategy and the Environment*, *24*(4), 252–265. https://doi.org/10.1002/bse.1821

Thoemmes, F., Rosseel, Y., & Textor, J. (2018). Local fit evaluation of structural equation models using graphical criteria. *Psychological Methods*, *23*(1), 27–41. https://doi.org/10.1037/met0000147

Thomson, S. B. (2010). *Sample Size and Grounded Theory* (SSRN Scholarly Paper 3037218). https://papers.ssrn.com/abstract=3037218

Tingley, D., Yamamoto, T., Hirose, K., Keele, L., & Imai, K. (2014). mediation: R Package for Causal Mediation Analysis. *Journal of Statistical Software*, *59*(5), 1–38. https://doi.org/10.18637/jss.v059.i05

Tong, X., & Su, J. (2018). Exploring young consumers' trust and purchase intention of organic cotton apparel. *Journal of Consumer Marketing*, *35*(5), 522–532. https://doi.org/10.1108/JCM-04-2017-2176

Touliatos, D., Dodd, I. C., & McAinsh, M. (2016). Vertical farming increases lettuce yield per unit area compared to conventional horizontal hydroponics. *Food and Energy Security*, *5*(3), 184–191. https://doi.org/10.1002/fes3.83

Tourangeau, R., Conrad, F. G., & Couper, M. P. (2013). Nonresponse in Web Surveys. In *The Science of Web Surveys*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199747047.003.0003

Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, *38*(1), 1–10. https://doi.org/10.1007/BF02291170

Tukker, A. (2015). Product services for a resource-efficient and circular economy – a review. *Journal of Cleaner Production*, *97*, 76–91. http://dx.doi.org/10.1016/j.jclepro.2013.11.049

Turker, D. (2009). Measuring Corporate Social Responsibility: A Scale Development Study. *Journal of Business Ethics*, *85*(4), 411–427. https://doi.org/10.1007/s10551-008-9780-6

Tymoshchuk, O., Lou, X., & Chi, T. (2024). Exploring Determinants of Second-Hand Apparel Purchase Intention and Word of Mouth: A Stimulus–Organism–Response Approach. *Sustainability*, *16*(11), Article 11. https://doi.org/10.3390/su16114445

UN Secretary-General, & World Commission on Environment and Development. (1987). *Report of the World Commission on Environment and Development: Note / by the Secretary-General.* https://digitallibrary.un.org/record/139811

Unipark (Version EFS 24.2). (2024). [Computer software]. Tivian XI GmbH. https://www.unipark.com/

Usrey, B., Palihawadana, D., Saridakis, C., & Theotokis, A. (2020). How Downplaying Product Greenness Affects Performance Evaluations: Examining the Effects of Implicit and Explicit Green Signals in Advertising. *Journal of Advertising*, *49*(2), 125–140. https://doi.org/10.1080/00913367.2020.1712274

Usunier, J.-C. (2006). Relevance in business research: The case of country-of-origin research in marketing. *European Management Review*, *3*(1), 60–73. https://doi.org/10.1057/palgrave.emr.1500049

Van Bussel, L. M., Kuijsten, A., Mars, M., & Van 'T Veer, P. (2022). Consumers' perceptions on food-related sustainability: A systematic review. *Journal of Cleaner Production*, *341*, 130904. https://doi.org/10.1016/j.jclepro.2022.130904

Van de Velde, L., Verbeke, W., Popp, M., Buysse, J., & Van Huylenbroeck, G. (2009). Perceived importance of fuel characteristics and its match with consumer beliefs about biofuels in Belgium. *Energy Policy*, *37*(8), 3183–3193. https://doi.org/10.1016/j.enpol.2009.04.022

Van der Linden, S. (2018). Warm glow is associated with low- but not high-cost sustainable behaviour. *Nature Sustainability*, *1*(1), 28–30. https://doi.org/10.1038/s41893-017-0001-0

Van der Velden, M. (2021). 'Fixing the World One Thing at a Time': Community repair and a sustainable circular economy. *Journal of Cleaner Production*, *304*, 127151. https://doi.org/10.1016/j.jclepro.2021.127151

Van Doorn, J., & Verhoef, P. C. (2011). Willingness to pay for organic products: Differences between virtue and vice foods. *International Journal of Research in Marketing*, *28*(3), 167–180. https://doi.org/10.1016/j.ijresmar.2011.02.005

Van Doorn, J., Verhoef, P. C., & Risselada, H. (2020). Sustainability Claims and Perceived Product Quality: The Moderating Role of Brand CSR. *Sustainability*, *12*(9), Article 9. https://doi.org/10.3390/su12093711

Villarejo-Ramos, A. F., & Sánchez-Franco, M. J. (2005). The impact of marketing communication and price promotion on brand equity. *Journal of Brand Management*, *12*(6), 431–444. https://doi.org/10.1057/palgrave.bm.2540238

Visser, M., Schoormans, J., & Vogtländer, J. (2018). Consumer buying behaviour of sustainable vacuum cleaners—Consequences for design and marketing. *Journal of Cleaner Production*, *195*, 664–673. https://doi.org/10.1016/j.jclepro.2018.05.114

Voorhees, C. M., Brady, M. K., Calantone, R., & Ramirez, E. (2016). Discriminant validity testing in marketing: An analysis, causes for concern, and proposed remedies. *Journal of the Academy of Marketing Science*, *44*(1), 119–134. https://doi.org/10.1007/s11747-015-0455-4

Waehning, N., & Filieri, R. (2021). Consumer motives for buying regional products: The REGIOSCALE. *Marketing Letters*, *33*, 215–236. https://doi.org/10.1007/s11002-021-09572-w

Wagner, M. M., & Heinzel, T. (2020). Human Perceptions of Recycled Textiles and Circular Fashion: A Systematic Literature Review. *Sustainability*, *12*(24), Article 24. https://doi.org/10.3390/su122410599

Wallach, K. A., & Popovich, D. (2023). When Big Is Less than Small: Why dominant brands lack authenticity in their sustainability initiatives. *Journal of Business Research*, *158*, 113694. https://doi.org/10.1016/j.jbusres.2023.113694

Wallner, T. S., Magnier, L., & Mugge, R. (2020). An Exploration of the Value of Timeless Design Styles for the Consumer Acceptance of Refurbished Products. *Sustainability*, *12*(3), Article 3. https://doi.org/10.3390/su12031213

Walsh, G., & Beatty, S. E. (2007). Customer-based corporate reputation of a service firm: Scale development and validation. *Journal of the Academy of Marketing Science*, *35*(1), 127–143. https://doi.org/10.1007/s11747-007-0015-7

Wang, F.-J., Hsiao, C.-H., Shih, W.-H., & Chiu, W. (2023). Impacts of Price and Quality Perceptions on Individuals' Intention to Participate in Marathon Events: Mediating

Role of Perceived Value. *Sage Open*, *13*(2). https://doi.org/10.1177/21582440231181431

Wang, H., Liu, H., Kim, S. J., & Kim, K. H. (2019). Sustainable fashion index model and its implication. *Journal of Business Research*, *99*, 430–437. https://doi.org/10.1016/j.jbusres.2017.12.027

Wang, P., & Kuah, A. T. H. (2018). Green marketing cradle-to-cradle: Remanufactured products in Asian markets. *Thunderbird International Business Review*, *60*(5), 783–795. https://doi.org/10.1002/tie.21925

Wang, W., Krishna, A., & McFerran, B. (2017). Turning off the Lights: Consumers' Environmental Efforts Depend on Visible Efforts of Firms. *Journal of Marketing Research*, *54*(3), 478–494. https://doi.org/10.1509/jmr.14.0441

Wang, X. (Shane), He, J., Curry, D. J., & Ryoo, J. H. (Joseph). (2022). Attribute Embedding: Learning Hierarchical Representations of Product Attributes from Consumer Reviews. *Journal of Marketing*, *86*(6), 155–175. https://doi.org/10.1177/00222429211047822

Wang, Y., & Hazen, B. T. (2016). Consumer product knowledge and intention to purchase remanufactured products. *International Journal of Production Economics*, *181*(Part B), 460–469. https://doi.org/10.1016/j.ijpe.2015.08.031i

Wang, Y., Hazen, B. T., & Mollenkopf, D. A. (2018). Consumer value considerations and adoption of remanufactured products in closed-loop supply chains. *Industrial Management & Data Systems*, *118*(2), 480–498. https://doi.org/10.1108/IMDS-10-2016-0437

Warren, C., Batra, R., Loureiro, S. M. C., & Bagozzi, R. P. (2019). Brand Coolness. *Journal of Marketing*, *83*(5), 36–56. https://doi.org/10.1177/0022242919857698

Wasaya, A., Saleem, M. A., Ahmad, J., Nazam, M., Khan, M. M. A., & Ishfaq, M. (2021). Impact of green trust and green perceived quality on green purchase intentions: A moderation study. *Environment, Development and Sustainability*, *23*(9), 13418–13435. https://doi.org/10.1007/s10668-020-01219-6

Webster, J., & Watson, R. T. (2002). Analyzing the Past to Prepare for the Future: Writing a Literature Review. *MIS Quarterly*, *26*(2), xiii–xxiii.

Webster, S., & Mitra, S. (2007). Competitive strategy in remanufacturing and the impact of take-back laws. *Journal of Operations Management*, *25*(6), 1123–1140. https://doi.org/10.1016/j.jom.2007.01.014

Weijters, B., & Baumgartner, H. (2012). Misresponse to Reversed and Negated Items in Surveys: A Review. *Journal of Marketing Research*, *49*(5), 737–747. https://doi.org/10.1509/jmr.11.0368

Weisstein, F. L., Meyer, J., & Kershaw, J. (2024). A matter of alignment? Effects of product types and environmental claim framing on consumer evaluation of sustainable foods. *Business Strategy and the Environment*, *33*(3), 1661–1674. https://doi.org/10.1002/bse.3565

Wendling, Z. A., Emerson, J. W., de Sherbinin, A., & Esty, D. C. (2020). *2020 Environmental Performance Index*. Yale Center for Environmental Law & Policy. https://epi.yale.edu/epi-results/2020/component/epi

White, H. (1980). A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. *Econometrica*, *48*(4), 817–838. https://doi.org/10.2307/1912934

White, K., Habib, R., & Hardisty, D. J. (2019). How to SHIFT Consumer Behaviors to be More Sustainable: A Literature Review and Guiding Framework. *Journal of Marketing*, *83*(3), 22–49. https://doi.org/10.1177/0022242919825649

White, K., Hardisty, D. J., & Habib, R. (2019, July 1). The Elusive Green Consumer. *Harvard Business Review*, *July–August 2019*. https://hbr.org/2019/07/the-elusive-green-consumer

Wickham, H., Averick, M., Bryan, J., Chang, W., D'Agostino McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., … Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686

Wiederhold, M., & Martinez, L. F. (2018). Ethical consumer behaviour in Germany: The attitude-behaviour gap in the green apparel industry. *International Journal of Consumer Studies*, *42*(4), 419–429. https://doi.org/10.1111/ijcs.12435

Wiley, J. F. (2023). *JWileymisc: Miscellaneous Utilities and Functions* (Version R package version 1.4.1) [Computer software]. https://CRAN.R-project.org/package=JWileymisc

Witten, I. H., Frank, E., Hall, M. A., & Pal, C. J. (2016). *Data Mining: Practical Machine Learning Tools and Techniques*. Morgan Kaufmann.

Wong, E. Y. C., Chan, F. F. Y., & So, S. (2020). Consumer perceptions on product carbon footprints and carbon labels of beverage merchandise in Hong Kong. *Journal of Cleaner Production*, *242*, 118404. https://doi.org/10.1016/j.jclepro.2019.118404

Wu, X., Lu, Y., Zhou, S., Chen, L., & Xu, B. (2016). Impact of climate change on human infectious diseases: Empirical evidence and human adaptation. *Environment International*, *86*, 14–23. https://doi.org/10.1016/j.envint.2015.09.007

Yang, Z., Sun, S., Lalwani, A. K., & Janakiraman, N. (2019). How Does Consumers' Local or Global Identity Influence Price–Perceived Quality Associations? The Role of Perceived Quality Variance. *Journal of Marketing*, *83*(3), 145–162. https://doi.org/10.1177/0022242918825269

Yap, B. W., & Sim, C. H. (2011). Comparisons of various types of normality tests. *Journal of Statistical Computation and Simulation*, *81*(12), 2141–2155. https://doi.org/10.1080/00949655.2010.520163

Ye, H., Bhatt, S., Deutsch, J., & Suri, R. (2022). Is there a market for upcycled pet food? *Journal of Cleaner Production*, *343*, 130960. https://doi.org/10.1016/j.jclepro.2022.130960

Zamani, B., Sandin, G., & Peters, G. M. (2017). Life cycle assessment of clothing libraries: Can collaborative consumption reduce the environmental impact of fast fashion? *Journal of Cleaner Production*, *162*, 1368–1375. https://doi.org/10.1016/j.jclepro.2017.06.128

Zander, K., & Hamm, U. (2010). Consumer preferences for additional ethical attributes of organic food. *Food Quality and Preference*, *21*(5), 495–503. https://doi.org/10.1016/j.foodqual.2010.01.006

Zeileis, A. (2004). Econometric Computing with HC and HAC Covariance Matrix Estimators. *Journal of Statistical Software*, *11*(10), 1–17. https://doi.org/doi.org/10.18637/jss.v011.i10

Zeileis, A., & Hothorn, T. (2002). Diagnostic Checking in Regression Relationships. *R News*, *2*(3), 7–10.

Zeileis, A., Köll, S., & Graham, N. (2020). Various Versatile Variances: An Object-Oriented Implementation of Clustered Covariances in R. *Journal of Statistical Software*, *95*(1), 1–36. https://doi.org/doi:10.18637/jss.v095.i01

Zeithaml, V. A. (1988). Consumer Perceptions of Price, Quality, and Value: A Means-End Model and Synthesis of Evidence. *Journal of Marketing*, *52*(3), 2–22. https://doi.org/10.1177/002224298805200302

Zeithaml, V. A., Jaworski, B. J., Kohli, A. K., Tuli, K. R., Ulaga, W., & Zaltman, G. (2020). A Theories-in-Use Approach to Building Marketing Theory. *Journal of Marketing*, *84*(1), 32–51. https://doi.org/10.1177/0022242919888477

Zeng, T., & Durif, F. (2020). The Impact of Eco-Design Packaging on Food Waste Avoidance: A Conceptual Framework. *Journal of Promotion Management*, *26*, 1–23. https://doi.org/10.1080/10496491.2020.1729320

Zheng, Q., Wang, H. H., & Lu, Y. (2018). Consumer Purchase Intentions for Sustainable Wild Salmon in the Chinese Market and Implications for Agribusiness Decisions. *Sustainability*, *10*(5), Article 5. https://doi.org/10.3390/su10051377

Zhou, H., Baskov, D., & Lease, M. (2013). Crowdsourcing Transcription Beyond Mechanical Turk. *Proceedings of the AAAI Conference on Human Computation and Crowdsourcing*, *1*(1). https://ojs.aaai.org/index.php/HCOMP/article/view/13093

Zimmer, L., & Swoboda, B. (2023). Perceived corporate social responsibility effects across nations – The role of national institutions. *International Business Review*, *32*(3), 102073. https://doi.org/10.1016/j.ibusrev.2022.102073

APPENDIX

Appendix A: Descriptions of products used in main quantitative study.

Chocolate bar:

Please imagine that you are shopping and notice a specific chocolate bar that is described as follows:

This chocolate is fully vegan and fair trade certified with cocoa beans sourced from South America and Africa. All ingredients are natural, organic, and clearly described on the packaging - which is made from a new type of paper. You can scan a QR code on the packaging to learn more about production and company principles.

Choose the flavor you like most, for example whole almond, salty caramel or dark (80%) chocolate.

You will be asked to evaluate the product described above in the following, please provide input to all questions. Your honest opinion is valued. There are no right or wrong answers.

Sweater:

Please imagine that you are shopping for a sweater and notice a specific one that is described as follows:

This sweater is made from 100% certified organic cotton. It does not contain any animal products and is approved by animal protection organizations. It has been manufactured in Europe under fair working conditions.

The sweater has a round neck, regular fit and can be machine-washed at 30 C°. Choose between several colors, for example beige, grey or black.

You will be asked to evaluate the product described above in the following, please provide input to all questions. Your honest opinion is valued. There are no right or wrong answers.

Appendix B: Items in the questionnaire in German.
Note: (*) = Reverse item.

Measure (source)	Items
PQ (Konuk, 2018; Stolz & Bautista, 2015)	1. Dieses Produkt ist an meine Bedürfnisse angepasst. 2. Dieses Produkt hat eine hohe Qualität. 3. Dieses Produkt ist hochwertig. 4. Dieses Produkt hat eine sehr gute Qualität.
PI (Das et al., 2020)	1. Ich werde dieses Produkt wahrscheinlich kaufen. 2. Ich bin gewillt, dieses Produkt zu kaufen. 3. Ich neige dazu, dieses Produkt zu kaufen.
PEF_uni (Magnier et al., 2016; Ye et al., 2022)	1. Dieses Produkt ist umweltfreundlich. 2. Dieses Produkt ist ein gutes Beispiel für ein umweltfreundliches Produkt. 3. Dieses Produkt ist nachhaltig.
Nat (Pula et al., 2014)	1. Es besteht aus natürlichen Inhaltsstoffen / Materialien. 2. Es wurde schonend verarbeitet. 3. Es ist frei von chemischen Substanzen. 4. Es enthält keine Zusatzstoffe. 5. Es enthält keine Rückstände von chemischen Sprays und Pestiziden.
Soc (Anagnostou et al., 2015)	1. Es wurde unter fairen Arbeitsbedingungen hergestellt. 2. Es ist ein sozial verantwortliches Produkt. 3. Es ist besser für das Wohl der Gesellschaft als andere Produkte. 4. Es wurde unter fairen Bedingungen gehandelt.
Ani (Lindeman & Väänänen, 2000; Stremmel et al., 2024)	1. Es wurde so produziert, dass Tiere keinen Schmerz erleiden mussten. 2. Es wurde so produziert, dass die Rechte von Tieren respektiert wurden. 3. Es wurde tierfreundlich hergestellt.
Cli (Stremmel et al., 2024)	1. Es ist vorteilhaft für das Klima im Vergleich zu ähnlichen Produkten. 2. Es ist gut für das Klima im Vergleich zu ähnlichen Produkten. 3. Es leistet einen Beitrag zum Klimaschutz. 4. Es verursacht weniger Treibhausgase (z.B. CO2) als ähnliche Produkte.
Res (Stremmel et al., 2024)	1. Es ist wassersparend. 2. Es ist mit schonender Ressourcennutzung verbunden. 3. Es nutzt Ressourcen effizient.
Dur (Hazen et al., 2017)	1. Defekte sind bei diesem Produkt selten. 2. Es hat eine angemessene Lebensdauer. 3. Es ist zuverlässig. 4. Es ist ein langlebiges Produkt.
Rep (Hazen et al., 2017)	1. Die Reparatur dieses Produkts sollte kein Problem sein. 2. Wenn es repariert werden muss, wird es nicht lange dauern. 3. Ich kann leicht Ersatzteile und Reparaturdienste für dieses Produkt finden. 4. Ich kann schnelle Reparaturdienste für dieses Produkt erhalten.
Rec	1. Es kann wiederverwendet werden.

Self-developed	2.	Es kann einfach recycelt werden.
	3.	Es erzeugt am Ende seiner Nutzungsdauer keinen unnötigen Abfall.
Wglow (Das et al., 2020)	1.	Ich fühle mich gut, wenn ich das Produkt kaufe, weil ich dabei helfen kann, die Umwelt zu schützen.
	2.	Durch den Kauf von diesem Produkt habe ich das Gefühl, dass ich zum Wohlergehen der Menschheit und Natur beitrage.
	3.	Durch den Kauf dieses Produkts fühle ich mich besser, weil es die Umwelt nicht schädigt.
Exper (Newell & Goldsmith, 2001)	1.	Der Hersteller hat viel Erfahrung mit umweltfreundlichen Produkten.
	2.	Der Hersteller ist versiert in der Herstellung von umweltfreundlichen Produkten.
	3.	Der Hersteller verfügt über umfangreiche Kenntnisse bei umweltfreundlichen Produkten.
	4.	Der Hersteller hat nicht viel Erfahrung mit umweltfreundlichen Produkten. (*)
Indic *Self-developed*	1.	Umweltfreundlichkeit und Qualität haben viel gemeinsam.
	2.	Meine Vorstellungen von Umweltfreundlichkeit und Qualität werden durch die gleichen Dinge beeinflusst (z.B. Gesundheit, Haltbarkeit, etc.).
	3.	Umweltfreundlichkeit und Qualität hängen zusammen.
Innov (Brockman & Morgan, 2003)	1.	Dieses Produkt fordert bestehende Ideen dieser Produktkategorie heraus.
	2.	Dieses Produkt bietet neue Ideen für diese Produktkategorie.
	3.	Dieses Produkt ist kreativ.
Effort (Bechwati & Xia, 2003)	1.	Hersteller investieren mehr Mühe in die Herstellung umweltfreundlicher Produkte als in konventionelle Produkte.
	2.	Hersteller betreiben keinen größeren Aufwand bei der Herstellung umweltfreundlicher Produkte als bei konventionellen Produkten. (*)
	3.	Hersteller arbeiten härter daran, umweltfreundliche Produkte herzustellen als konventionelle Produkte.
Stren (Mai et al., 2019)	1.	Konventionelle Produkte sind oft effektiver als umweltfreundlich hergestellte Produkte.
	2.	Umweltfreundliche Produkte sind weniger stark.
	3.	Reguläre Produkte sind oft stärker als die umweltfreundlichen Alternativen.
	4.	Eco Umweltfreundliche Produkte sind weniger haltbar, weniger robust, weniger geschmackvoll, etc.
	5.	Herkömmliche Produkte sind oft haltbarer, robuster, etc. als umweltfreundlich hergestellte Produkte.

Appendix C: Internal consistency measures based on CFA with Chocolate and Sweater combined.

Factor	Alpha	Omega	AVE	Items	Factor loading
PQ	**0.90**	**0.90**	**0.75**	1. -	
				2. This product is of high quality.	0.911
				3. This product is a high standard product.	0.800
				4. This product is of very good quality.	0.887
PI	**0.93**	**0.93**	**0.81**	1. I am likely to buy this product.	0.907
				2. I am willing to buy this product.	0.911
				3. I am inclined to buy this product.	0.886
PEF_uni	**0.91**	**0.91**	**0.77**	1. This product is an eco-friendly product.	0.908
				2. This product is a good example of an eco-friendly product.	0.898
				3. This product is sustainable.	0.826
Nat	**0.75**	**0.77**	**0.46**	1. It contains natural ingredients / materials.	0.578
				2. -	
				3. It is free of chemical substances.	0.764
				4. It contains no additives.	0.629
				5. It is free of residues from chemical sprays and pesticides.	0.681
Soc	**0.83**	**0.84**	**0.63**	1. It has been created under fair working conditions.	0.857
				2. It is a socially responsible product.	0.721
				3. -	
				4. It has been traded under fair conditions.	0.797
Ani	**0.90**	**0.90**	**0.75**	1. It has been produced in a way that animals have not experienced pain.	0.834
				2. It has been produced in a way that animals' rights have been respected.	0.884
				3. It has been produced considering animal welfare.	0.877
Cli	**0.86**	**0.86**	**0.60**	1. It is beneficial for the climate compared to similar products.	0.836
				2. It is good for the climate compared to similar products.	0.808
				3. It contributes to climate protection.	0.774
				4. It causes less greenhouse gases (e.g., CO2) than similar products	0.702
Res	**0.72**	**0.74**	**0.48**	1. It is water-saving.	0.547
				2. Is associated with low-impact land use.	0.778
				3. It is resource-efficient.	0.767
Dur	**0.81**	**0.82**	**0.53**	1. Failure of this product does not occur often.	0.593
				2. It has an adequate lifespan.	0.800
				3. It is reliable.	0.704
				4. It is a durable product.	0.789
Rep	**0.93**	**0.93**	**0.78**	1. Getting it repaired should not be an issue.	0.898

				2. When it needs repair, it will not take long.	0.872
				3. I can receive responsive repair services for this product.	0.885
				4. I can easily find repair parts and services for this product.	0.878
Wglow	**0.88**	**0.89**	**0.72**	1. I will feel good by purchasing this product because I will be able to help in protecting the environment.	0.877
				2. With the purchase of this product, I will have the feeling of contributing to the wellbeing of humanity and nature.	0.849
				3. With the purchase of this product, I will feel better because it does not harm the environment.	0.818
Innov	**0.84**	**0.84**	**0.65**	1. This product challenges existing ideas for this product category.	0.766
				2. This product offers new ideas to this product category.	0.887
				3. This product is creative.	0.745
Exper	**0.89**	**0.89**	**0.74**	1. Its producer has a great amount of experience with eco-friendly products.	0.856
				2. Its producer is skilled with eco-friendly products.	0.842
				3. Its producer has great expertise with eco-friendly products.	0.880
				4. -	
Indic	**0.77**	**0.81**	**0.60**	1. Eco-friendliness and quality have a lot in common.	0.890
				2. My perceptions of eco-friendliness and quality are influenced by the same things (e.g., healthiness, durability, etc.).	0.506
				3. Eco-friendliness and quality are related.	0.827
Effort	**0.71**	**0.72**	**0.48**	1. Producers put more effort into creating eco-friendly products than conventional products.	0.810
				2. Producers do not exert more effort in creating eco-friendly products than conventional products. (*)	0.493
				3. Producers work harder to create eco-friendly products than conventional products.	0.721
Stren	**0.88**	**0.88**	**0.60**	1. Conventional products are often more effective than eco-friendly produced products.	0.701
				2. Eco-friendly products are less strong.	0.780
				3. Regular products often are stronger than the eco-friendly alternatives.	0.770
				4. Eco-friendly products are less durable, less robust, less tasty, etc.	0.799
				5. Conventional products are often more durable, robust, etc., than products that are manufactured eco-friendly.	0.805

Appendix D: Comparison of regression results with in- or excluded InfObs for Chocolate. Explanation of models: (1) all observations, (2) without InfObs identified controlling for age, gender, familiarity and country, (3) without InfObs identified without considering control variables, and (4) without InfObs identified from individual regressions.

	Dependent variable: PQ			
	(1)	**(2)**	**(3)**	**(4)**
	β (SE)	β (SE)	β (SE)	β (SE)
PEF_dim	0.235***	0.275***	0.281***	0.274***
	(0.061)	(0.058)	(0.058)	(0.059)
Wglow	0.119*	0.173***	0.132*	0.161**
	(0.070)	(0.067)	(0.070)	(0.064)
Innov	0.002	-0.008	0.014	0.002
	(0.059)	(0.052)	(0.055)	(0.052)
Stren	-0.093**	-0.090**	-0.101**	-0.099**
	(0.046)	(0.045)	(0.047)	(0.045)
Indic	0.024	0.021	0.002	0.021
	(0.046)	(0.047)	(0.050)	(0.047)
Effort	0.102**	0.082*	0.103**	0.095**
	(0.044)	(0.044)	(0.047)	(0.043)
Exper	0.235***	0.195***	0.256***	0.193***
	(0.059)	(0.058)	(0.064)	(0.055)
Age	-0.011***	-0.010***	-0.007**	-0.008**
	(0.004)	(0.004)	(0.003)	(0.003)
Male	-0.152*	-0.181**	-0.204**	-0.203**
	(0.087)	(0.087)	(0.083)	(0.082)
Familiar	0.088	0.090*	0.089	0.076
	(0.057)	(0.053)	(0.055)	(0.054)
Country_other	-0.182*	-0.175	-0.203*	-0.163
	(0.109)	(0.110)	(0.105)	(0.104)
PEF_dim:Stren	-0.039	-0.022	0.009	0.014
	(0.060)	(0.049)	(0.049)	(0.049)
PEF_dim:Indic	0.018	-0.013	-0.016	-0.019
	(0.056)	(0.047)	(0.050)	(0.046)
Constant	0.088	0.073	-0.017	0.076
	(0.276)	(0.259)	(0.264)	(0.255)
Observations	367	351	344	352
R^2	0.369	0.398	0.404	0.368
Adjusted R^2	0.345	0.374	0.380	0.344
Residual Std. Error	0.811 (df = 353)	0.764 (df = 337)	0.762 (df = 330)	0.763 (df = 338)
F Statistic	15.846*** (df = 13; 353)	17.104*** (df = 13; 337)	17.201*** (df = 13; 330)	15.158*** (df = 13; 338)

*p < 0.1; **p < 0.05; ***p < 0.01

Appendix E: Comparison of regression results with in- or excluded InfObs for Sweater. Explanation of models: (1) all observations, (2) without InfObs identified controlling for age, gender, familiarity and country, (3) without InfObs identified without considering control variables, and (4) without InfObs identified from individual regressions.

	Dependent variable: PQ			
	(1)	**(2)**	**(3)**	**(4)**
	β (SE)	β (SE)	β (SE)	β (SE)
PEF_dim	0.322***	0.346***	0.352***	0.323***
	(0.059)	(0.062)	(0.063)	(0.069)
Wglow	0.186***	0.227***	0.246***	0.154**
	(0.068)	(0.067)	(0.076)	(0.068)
Innov	0.038	0.008	-0.024	0.030
	(0.059)	(0.056)	(0.055)	(0.060)
Stren	-0.084*	-0.048	-0.063	-0.042
	(0.050)	(0.049)	(0.051)	(0.049)
Indic	0.135**	0.155***	0.142***	0.129**
	(0.055)	(0.058)	(0.057)	(0.056)
Effort	0.001	-0.024	-0.033	-0.002
	(0.046)	(0.046)	(0.048)	(0.046)
Exper	-0.066	-0.072	-0.070	-0.044
	(0.057)	(0.059)	(0.060)	(0.057)
Age	-0.011**	-0.009**	-0.007	-0.008**
	(0.004)	(0.004)	(0.004)	(0.004)
Male	-0.081	-0.069	-0.057	-0.040
	(0.090)	(0.089)	(0.088)	(0.089)
Familiar	0.032	0.028	0.045	0.044
	(0.056)	(0.056)	(0.054)	(0.056)
Country_other	-0.119	-0.148	-0.118	-0.176
	(0.124)	(0.129)	(0.127)	(0.122)
PEF_dim:Stren	0.005	-0.002	0.007	0.0001
	(0.048)	(0.053)	(0.057)	(0.055)
PEF_dim:Indic	-0.029	-0.066	-0.028	0.023
	(0.045)	(0.054)	(0.051)	(0.057)
Constant	0.299	0.304	0.098	0.164
	(0.289)	(0.289)	(0.281)	(0.287)
Observations	365	351	341	350
R^2	0.317	0.354	0.347	0.275
Adjusted R^2	0.292	0.329	0.321	0.247
Residual Std. Error	0.843 (df = 351)	0.797 (df = 337)	0.787 (df = 327)	0.817 (df = 336)
F Statistic	12.542*** (df = 13; 351)	14.207*** (df = 13; 337)	13.342*** (df = 13; 327)	9.826*** (df = 13; 336)

*p < 0.1; **p < 0.05; ***p < 0.01

Appendix F: Results for the analysis of predictive value – Chocolate (PEF_uni).

PLS in sample metrics:

	PQ2	PQ3	PQ4	PI1	PI2	PI3
RMSE	0.727	0.789	0.758	1.207	1.186	1.231
MAE	0.587	0.621	0.605	0.932	0.886	0.939

PLS out of sample metrics:

	PQ2	PQ3	PQ4	PI1	PI2	PI3
RMSE	0.759	0.816	0.792	1.221	1.201	1.246
MAE	0.610	0.643	0.628	0.940	0.896	0.950

MLR in sample metrics:

	PQ2	PQ3	PQ4	PI1	PI2	PI3
RMSE	0.683	0.746	0.712	1.090	1.089	1.142
MAE	0.554	0.594	0.569	0.865	0.825	0.890

MLR out of sample metrics:

	PQ2	PQ3	PQ4	PI1	PI2	PI3
RMSE	0.763	0.842	0.800	1.229	1.223	1.273
MAE	0.618	0.661	0.632	0.965	0.918	0.984

Appendix G: Results for the analysis of predictive value – Sweater (PEF_uni).

PLS in sample metrics:

	PQ2	PQ3	PQ4	PI1	PI2	PI3
RMSE	0.798	0.848	0.811	1.026	0.945	1.035
MAE	0.638	0.679	0.631	0.809	0.752	0.789

PLS out of sample metrics:

	PQ2	PQ3	PQ4	PI1	PI2	PI3
RMSE	0.828	0.880	0.840	1.036	0.955	1.044
MAE	0.664	0.705	0.655	0.817	0.758	0.793

MLR in sample metrics:

	PQ2	PQ3	PQ4	PI1	PI2	PI3
RMSE	0.743	0.761	0.744	0.949	0.877	0.971
MAE	0.593	0.603	0.584	0.753	0.692	0.748

MLR out of sample metrics:

	PQ2	PQ3	PQ4	PI1	PI2	PI3
RMSE	0.830	0.856	0.846	1.054	0.988	1.085
MAE	0.659	0.674	0.649	0.833	0.772	0.827