ESSAYS ON CONSUMER PERCEPTION
APPLICATIONS TO INFLATION, BUSINESS-TO-BUSINESS BRANDS,
AND RESPONSE STYLES

Inauguraldissertation

zur

Erlangung des Doktorgrades

der

Wirtschafts- und Sozialwissenschaftlichen Fakultät

der

Universität zu Köln

2016

vorgelegt von

Diplom-Wirtschafts-Ingenieur Eric Michael Lennartz

aus

Düren

Referent: Prof. Dr. Marc Fischer

Korreferent: Prof. Dr. Hernán A. Bruno

Tag der Promotion: 09. Februar 2017

AKNOWLEDGMENTS

Contents

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

#	Number of
ARS	Acquiescence Response Style
B2B	Business-to-business
B2C	Business-to-consumer
COICOP	Classification of Individual Consumption According to Purpose
DARS	Disacquiescence Response Style
d.f.	degrees of freedom
ERS	Extreme Response Style
ERT	Egoistic Response Tendency
EU	European Union
FEB	Frequency Estimation Bias
IO	Item Omission
IT	Information Technology
MAE	Mean Absolute Error
MRS	Midpoint Response Style
MRT	Moralistic Response Tendency
MSE	Mean Squared Error
N	Sample Size
NCR	Noncontingent Response Style
RR	Response Range
S	Speed
SD	Standard Deviation
SDR	Social Desirable Responding
TL	Tendency to the Left
TR	Tendency to the Right

UK	United Kingdom of Great Britain and Northern Ireland
US	United States of America
USA	United States of America

LIST OF SYMBOLS

PROJECT 1

Variables

AADC	Availability across Distribution Channels
AOS	Affinity for Online Shopping
BL	Brand Loyalty
CC	Category Clutter
CE	Customer Expertise
DPC	Direction of Price Change
DUR	Dummy Variable Indicating that a Product Class Contains Mainly Durables
ES	Economic Skills
FRE	Purchase Frequency
MI	Monetary Impact
MS	Mathematical Skills
NI	Net Income
NPC	Number of Product Classes
PC	Price Consciousness
PCR	Price Change Range
PI	Perceived Inflation (Scale Value)
PPPF	Perceived Price Promotion Frequency
PQR	Price-quality Relationship
REC	Purchase Recency
SA	Level of Social Amplification
SER	Dummy Variable Indicating that a Product Class Contains Mainly Services

SEX	Dummy Variable Indicating that a Person is Male
SOCIODEMOGRAPHICS	Vector Containing the Gender (male = 1), Age, Income, Income Change (Decrease = -1, Stable = 0, Increase = 1), Individuals' Involvement in Daily Shopping Activities, Household Size, Education (Graduate = 1), Profession (Student = 1), and Marital Status (Living in a Partnership = 1)
TSQ	Time Spent for the Questionnaire
π	Inflation Rate (Percentage Value)
π^p	Perceived Inflation (Percentage Value)

Indexes

i	Product Group Index with i = 1,… *I* (Number of Product Groups)
j	Product Class Index with j = 1,… *J* (Number of Product Classes in a product group)
k	Respondent Index with k = 1,… *K* (Number of Respondents)
s	Stimulus Index with s = 1,… *S* (Number of Stimuli)

Model Parameters

α, β, γ	Regression Parameters to Be Estimated
ε, ζ, ν	Error Terms
Σ	Variance-covariance Matrix of Random Parameters
w	Weighting Parameter for Weighted Models

PROJECT 2

Variables

BR	Business Reach
BS	Brand Strength
C	Communication Performance
D	Distribution Performance
D_{COU}	Dummy Variable for the Country
D_{IND}	Dummy Variable for the Industry
FS	Firm Size
IE	Innovation and Expertise
PD	Product Performance
PR	Price Performance
ROLE	Role of Respondent (Taking Action within the Decision Process within a Group = 0, Taking Action within the Decision Process Alone = 1)
SC	Sustainability and Corporate Governance
SR	Share of Revenue from B2C Compared to B2B Activities

Indexes

j	Index for Observation j = 1,… *J* (Number of Observations)
k	Index for Respondent k = 1,… *K* (Number of Respondents)
l	Index for Country l = 1,… L (Number of Countries)
s	Index for Industry s = 1,… *S* (Number of Industries)

Model Parameters

β, γ, δ	Regression Parameters to Be Estimated
ε	Error Term

PROJECT 3

Variables

BRANDRELEVANCE	B2B Brand Relevance
BUSSINESSREACH	Global Business Reach of Respondents' Company (Globally = 1, Home Continent = 2, Home Country = 3, Specific Region of Our Home Country = 4)
BUYINGCENTERSIZE	Buying Center Size in Respondents' Company (Only Me = 1, Me and Another Person = 2, Me and Two Other Persons = 3, Me and Three Other Persons = 4, Me and More than Three Other Persons = 5)
HETEROGENEITY	Provider Heterogeneity
IMAGE	Importance of Brand's Imagery Benefit
VISIBILITY	Visibility of a B2B Branded Product to End Customers
REPORTINGLEVELS	Number of Reporting Levels to the CEO (I Am the CEO = 0, I Directly Report to the CEO = 1, 2 Levels = 2, 3 Levels =3, 4 Levels = 4, 5 or More Levels = 5)
RESPONSIBILITY	Responsibility of Respondent in B2B Purchase Decisions (Taking Action within the Decision Process within a Group = 0, Taking Action within the Decision Process Alone = 1)
RISK	Importance of Brand's Risk Reduction Function
SIZE	Number of Employees of Respondents' Company (100-249 = 1, 250-999 = 4, 1,000-4,999=20, 5,000-9,999=40, 10,000 and More = 60)

Indexes

i	Index for Respondent i = 1,… *I* (Number of Respondents)
j	Index for Industry j = 1,… *J* (Number of Industries)
k	Index for Category that Respondents Evaluate k = 1,… *K* (Number of Categories)
m	Index for country m = Germany, USA
n	Index for Model Parameter n = 0 (Intercept), 1 (Risk Reduction Function, 2 (Imagery Benefit Function)

Model Parameters

β, γ, λ	Regression Parameters to Be Estimated
$\vartheta, \eta, \zeta, \xi$	Error Terms
$\mathbf{T}$	Variance-covariance Matrix of Random Parameters

1 Overview

The objective of this cumulative dissertation is to gather new insights on consumers' perception processes and their measurement. It consists of four projects. Table 1 provides an overview on the projects of this dissertation and shows their co-authors and status in the publication process.

Table 1: Overview of Dissertation Projects

Project	Title	Author(s)	Status
1	Consumers' Perception of Price Inflation	Eric Michael Lennartz and Marc Fischer	Prepared to submit to: *Journal of Consumer Research*
2	Drivers of B2B Brand Strength – Insights from an International Study across Industries	Eric Michael Lennartz, Marc Fischer, Manfred Krafft, and Kay Peters	Published in: *Schmalenbach Business Review*
3	The Relevance of Brands in B2B Markets – An International Study across Categories	Eric Michael Lennartz and Marc Fischer	Prepared to submit to: *Journal of the Academy of Marketing Science*
4	Culture and Survey Research: A Review on How Response Styles Differ across Cultures	Eric Michael Lennartz	Prepared to submit to: *Psychology and Marketing*

The first project of this thesis addresses how consumers perceive price inflation and how researchers can measure this perceived inflation. However, the term consumer is not limited to the standard consumer as a person on a B2C (business-to-consumer) market, but also refers to organizations as consumers on B2B (business-to-business) markets. Thus, the second and third project of this dissertation analyze how managers and organizational buyers perceive brands. In detail, the second project focuses on when and how managers perceive B2B brands to be strong, while the third project assesses when and how managers perceive B2B brands to be relevant for their purchase decisions. While the first three projects of this thesis build on surveys to measure latent constructs, the fourth project deals with the perception and validity

of survey measurements between cultures. It reviews how culture shapes response styles to survey items.

Therefore, this dissertation contributes to three well-known research gaps: (1) How does subjective data help explain economic phenomena (Katona 1974), (2) how do "organizational buyers make product choice decisions" (Brown et al. 2011, p. 202), and (3) how do "population differences in cognition and communication affect the response process" (Schwarz 2003, p. 592).

2 Introduction

In everyday business, marketing works along the chain of stimulus, organism including perception, and response (Woodworth 1929). Marketers set stimuli, consumers perceive and evaluate these stimuli, and, finally, might choose a response.

Stimuli set by marketers normally consider the marketing mix and its dimensions. However, stimuli may also come from consumers' environment and other cues. Altogether, consumers experience these stimuli consciously or nonconsciously. Accordingly, they form evaluations and associations such as a brand's image during their perception processes. This perception may be complex and involve various stages. Particularly, it is prone to individual moderators, like a consumer's demographics, psychographics, culture, or organizational cues which may influence how s/he actually perceives a stimulus. Depending on these perceptions, consumers react, for example, by purchases or recommendations, but may also choose to show no action. Figure 1 shows and sorts the projects of this thesis along this overall process.

Figure 1: Classification of Dissertation Papers along the Stimulus-Organism-Response Model

	Stimulus	▶	Organism: Consumer Perception	▶	Response

Project 1

	Product Price Changes	▶	Perceived Inflation	▶	Consumer Behavior

Individual Moderators
- Product Characteristics
- Consumer Psychographics
- Consumer Demographics

Project 2

	B2B Marketing Mix B2B Brand Associations	▶	B2B Brand Strength	▶	Purchase Decision

Individual Moderators
- Manager Characteristics
- Organizational Cues
- Product Category
- Country

Project 3

	B2B Brands	▶	B2B Brand Relevance	▶	Purchase Decision

Individual Moderators
- Manager Characteristics
- Organizational Cues
- Product Category
- Country

Project 4

	Questionnaire Item	▶	Item Comprehension, Retrieval, and Judgment	▶	Item Response/ Response Style

Individual Moderators
- Culture
- Acculturation

Shaded parts are not focus of the project

One important objective of marketing research is to uncover perception processes between the implementation of a marketing activity and the actual customer reaction. Only if marketing managers know how consumers perceive marketing activities, they are able to implement their marketing activities effectively. However, in many areas research still lacks a sufficient analysis of perception processes.

In detail, a research gap exists concerning the perception of price inflation. Hereby, consumers perceive stimuli such as general price movements. Marketers may only partly influence these environmental cues because they involve external sources, such as competitors or media. Nevertheless, these cues appeal as stimuli to the consumers and are the basis for consumers' perceptions of price inflation. Thus, like marketing activities, these cues extensively drive consumer reactions and behavior. For example, consumers may reduce their consumption volume in times of perceived inflation (Katona 1974). Therefore, project one "Consumers' Perception of Price Inflation" (by Eric Michael Lennartz and Marc Fischer) addresses how consumers perceive price inflation across product groups depending on consumer demographics and psychographics and how researchers and managers may measure this perceived price inflation. Thus, we provide researchers with valuable insights on price perception and corresponding consumer behavior and a tool to model perceived inflation in future research projects.

Another area that lacks research is the perception of B2B brands (Brown et al. 2011). Despite the limited research, B2B brands have substantive relevance for marketing management and research. They have the potential to sustainably enhance company revenues (Fischer, Giehl, and Freundt 2011). Hence, project two "Drivers of B2B Brand Strength – Insights from an International Study across Industries" (by Eric Michael Lennartz, Marc Fischer, Manfred Krafft, and Kay Pieters) and three "The Relevance of Brands in B2B Markets – An International Study across Categories" (by Eric Michael Lennartz and Marc Fischer) address this

topic. While project two outlines when B2B consumers perceive brands to be strong, project three assesses when B2B consumers perceive brands as relevant for purchase decisions. For both projects applies, notwithstanding individual consumer perception, that, at the organism stage, business scenarios involve two levels of perception. Like on B2C markets, perception applies to a first individual perspective, namely the manager perspective. However, on a second level, it involves the organizational perspective which does not exist on B2C markets. Accordingly, both projects deepen the knowledge on how organizational factors influence perception processes. Moreover, both projects assess phenomena across B2B product categories in three (project 2) or two countries (project 3). Thus, they also provide insights on how B2B branding processes differ across product categories and countries.

One obstacle to the research of perception processes is the fact that these processes take place in the mind of the consumer and are, thus, not directly observable from the outside. Researchers use surveys to uncover these latent processes. Therefore, all projects of this dissertation involve the measurement via latent constructs in surveys. Similar to the perception processes above, survey response fits to the stimulus-organism-response scheme. Here, items are the stimuli, the interpretation, judgment and retrieval phase overtake the organism phase, including respondents' perception, and survey response equals the behavioral response (Tourangeau, Rips, and Rasinski 2009).

Importantly, during the perception and response phase, surveys entail the risk that respondents systematically deviate in their answers from their actual perception (Baumgartner and Steenkamp 2001). Such deviations are called response styles. These styles extensively depend on individual consumer characteristics. Among others, culture may especially shape these response styles (e.g., Wong, Rindfleisch, and Burroughs 2003). The fourth project of this dissertation "Culture and Survey Research: A review of How Response Styles Differ across Cultures" (by Eric M. Lennartz), thus, reviews studies that address how and when a

respondent's cultural background promotes a systematic bias in her/his response to a survey item. Thereby, the cultural background of an individual does not only cover her/his initial beliefs, values, and norms, but also those beliefs, values, and norms learned from acculturation processes. In detail, acculturation describes the process that occurs, if individuals with an initial cultural background enter an environment where another culture is prevalent. They learn about the new culture, decide whether and how to adapt to it and apply strategies to avoid conflicts between both cultures (e.g., Ownbey and Horridge, 1997; Peñaloza, 1994). Accordingly, they may develop new cultural patterns.

3   Summary of Dissertation Projects

3.1 Project 1: Consumers' Perception of Price Inflation

While statistical offices put strenuous efforts into the measurement of price inflation, consumers' "perceived inflation" deviates strongly from the official figures (e.g., Bates and Gabor 1986). Nonetheless, perceived inflation has a strong impact on research and management. It may shape price fairness perceptions (Homburg, Hoyer, and Koschate 2005), increase or decrease price sensitivity (Bijmolt, van Heerde, and Pieters 2005), and cause shifts in consumption behavior and volumes (Katona 1974). Notwithstanding, neither a valid measurement, nor a conceptualization of perceived inflation exists that suits marketing purposes.

The objective of this project is, hence, to develop a psychological rating scale for the measurement of perceived inflation. Based on literature, discussions with experts and potential respondents, and several pre-tests we develop a three-item psychological scale to measure perceived inflation. Final items reflect three different ways of how companies change prices, namely to changing nominal prices, quantities, or quality of products and services. We apply this scale in a first consumer survey to 13 product groups and 53 corresponding product classes. In addition, we gather data on validity measures and potential drivers of perceived infla-

6

tion. Our scale shows excellent reliability. Based on validity measures, we also find good convergent, nomological, and divergent validity.

As rating values of perceived inflation on our scale are not directly comparable to the percentage values of official inflation, we conduct a second consumer survey. The objective of this survey is to determine a mapping function that maps rating scale values onto percentage rates. Therefore, the second survey comprises a quasi-experiment. Here, we provide respondents with several stimuli that each contain one price change. Respondents rate these price changes on our rating scale. We estimate a mapping function that regresses these rating values on percentage values of the given price changes. We use this function to transform our original scale values to individual percentage values for each participant of our first survey.

For some product groups, perceived inflation deviates strongly to the real inflation rates. For example, consumers perceive an inflation rate for transportation services that exceeds the actual inflation rate by 21.42 percent points. Likewise, the average perceived inflation rate for newspapers, books, and stationery, is 5.47 percent points lower than the official inflation rate for this product group.

Two factors seem to drive these deviations. On the one hand, perceived inflation for the underlying product classes differs from the official rates. On the other hand, consumers apply a subjective weighting scheme (perceived market basket) when they aggregate product class inflation rates for product groups and higher aggregation levels. This weighting strongly differs from the official market basket that involves average spending across all consumers per product class. Results, thus, indicate that consumers actually involve two steps when perceiving inflation. They subjectively perceive price trends for products and product classes and then weight them subjectively to form inflation rates for higher aggregation levels.

The extent of these two deviations particularly varies with consumer and product characteristics. Therefore, we conduct a driver analysis to test which parameters drive the deviation

of perceived and official inflation for product classes and the weights in consumers' perceived market baskets. Importantly, we find that drivers may show complex effects because they may individually influence both processes. For example, deviations between perceived and official inflation are higher for product classes with more frequent and recent purchases, but consumers also assign higher weights to these classes in their perceived market baskets. Thus, deviations between perceived and official inflation may be even more pronounced for higher aggregation levels.

Overall, we provide researchers, managers, and economists with a tool that allows them to track perceived inflation. They can use these values to gain more precise insights compared to, if they used actual inflation rates. Moreover, our driver analysis offers insights on which product classes may especially drive the deviation of perceived and official inflation. Moreover, managers should track classes with a frequent purchase pattern because they are likely to deviate in their inflation perceptions and consumers assign a high weight to them. These classes act as indicators for future consumption shifts.

3.2 Project 2: Drivers of B2B Brand Strength – Insights from an International Study across Industries

Research concerning brand management almost exclusively focuses on B2C brands and leaves out B2B brands (Homburg, Klarmann, and Schmitt 2010; Sethuraman, Tellis, and Briesch 2011). Considering actual brand performance, this lack in research appeals surprising. In 2015, three primarily B2B brands, namely General Electric, IBM, and Microsoft, have ranked among the ten strongest brands in the Interbrand (2016) brand index. Nevertheless, research on B2B brand management is still scarce. Importantly, a research gap exists on the drivers of B2B brand strength. We, thus, lack knowledge on how to build up strong B2B brands. However, this knowledge is pivotal. It enables companies to develop brand profiles and strategies to build up strong brands and to make profitable brand investments.

Likewise, we cannot simply transfer results that researchers have found for B2C brands to B2B brands. This approach bears severe shortcomings because influences on B2B brands differ substantially from influences on B2C brands (e.g., Homburg, Klarmann, and Schmitt 2010). For example, group dynamics may be more pronounced in B2B purchases because buying centers make decisions here. Thus, the first objective of this project is to develop a B2B-specific concept of brand strength. Based on this concept, we conduct a driver analysis that shows which factors promote B2B brand strength.

For the concept of B2B brand strength, we expect that the B2B marketing mix and brand associations are major drivers of brand strength. In detail, associations should reflect a company's abilities and social responsibility (Brown and Dacin 1997). To test this framework, we conduct a survey for B2B brands in seven industries. Participants are B2B managers from Germany, India, and the US. For each manager, we collect data on B2B brand strength of actual providers in an industry and potential drivers of this brand strength. Especially, we include items that measure potential brand associations from our concept of B2B brand strength.

In a first analysis, we conduct an exploratory factor analysis for these association items. Two factors evolve that summarize managers' B2B brand associations. These factors are "sustainability and corporate governance" and "innovation and expertise". As expected, they split up associations into a company's abilities and social responsibility.

In a second step, we include these factors as associations in our driver analysis. We assess drivers on an overall level and separately for each country and industry. In addition to the two brand association factors, we include the marketing mix perception of B2B providers in the analysis. The results show that the brand associations "sustainability and corporate governance" and "innovation and expertise" are main drivers of B2B brand strength. Nevertheless, their influence differs across countries and industries. For example, "innovation and ex-

pertise" shows an effect on brand strength for all industries except the finance and insurance industry.

Concerning the marketing mix, we find that managers' perception of product and distribution are additional major drivers, while price and communication show weaker effects. This pattern is even more pronounced across countries and industries. Communication shows no significant effect in Germany and India, and, on the industry level, it is only highly significant for the finance and insurance industry. Therefore, traditional communication instruments seem to be far less effective to create brand strength than in B2C markets. Instead, companies should focus on the touchpoints between employees and B2B consumers. As the effects of product and distribution show, they are most influential for B2B brand strength.

However, companies should not solely focus on product and distribution performance. Our model shows that brand associations as intangible components are highly relevant for brand strength. Therefore, companies succeed which understand to combine both aspects. They need to fulfill customer requirements for product and distribution and link their brands to the associations "sustainability and corporate governance" and "innovation and expertise".

3.3 Project 3: The Relevance of Brands in B2B Markets – An International Study across Categories

In B2B markets, critics argue that managers decide rationally and, thus, B2B brands have no relevance (Brown et al. 2011). Nevertheless, studies show that, if companies build up B2B brands, they may strongly increase their profits (Fischer, Giehl, and Freundt 2011). B2B managers, indeed, consider brands in their purchase decisions and, thus, behave non-rational concerning brands. Meanwhile, decision criteria and purchase processes differ from the equivalents on B2C markets (Zablah, Brown, and Donthu 2010). For example, decision processes for B2B purchases are often institutionalized and involve a buying center. Results and concepts that hold for B2C brands, are, hence, not directly transferable to B2B brand manage-

ment. Therefore, research lacks insights on when and to which extent B2B consumers consider brands in their purchase decisions. Accordingly, this project assesses how and when brands are relevant in B2B markets.

As on B2C markets (Fischer, Völckner, and Sattler 2010), brands should not be equally relevant across B2B product categories. This issue is of particular importance because brand relevance is necessary to generate profits from favorable brand associations and brand knowledge (Keller 1993). If brands are highly relevant for B2B customers, companies can use and create strong brands to gain additional profits. In contrast, if brands are not relevant for B2B customers at all, even the strongest brand does not work as a competitive advantage in the short-term. Nonetheless, as brand relevance is not constant over time, companies may use long-lasting brand investments to create brand relevance in a category. This notion does not only hold for B2C brands, but is also true for B2B brands. The knowledge of brand relevance in a category or specific purchase situations, thus, carries ample insights for companies' future brand investments and selling strategies. The objective of this project is, hence, to measure B2B brand relevance across product categories and countries and to show which factors drive this relevance.

Therefore, we first develop a concept of B2B brand relevance that relates brand relevance to B2B-specific brand functions. These brand functions differ from those functions on B2C markets. On B2B markets, managers primarily use brands to bundle information. Differently to B2C markets, this information efficiency function is a second order construct that is reflected by two subfunctions. Brands bundle information on the risk of a purchase decision and the imagery benefits of a brand. Again, these subfunctions have a broader scope than seemingly equivalents on B2B markets, such as the risk reduction and social demonstration function by Fischer, Völckner, and Sattler (2010). For example, the B2B risk reduction function

involves a personal and organizational risk component, while the B2C risk reduction function only aims at the personal risk.

To validate this concept, we conduct a survey among B2B managers from Germany and the US. Among other scales, it contains scales for B2B brand relevance and potential brand functions. We develop these scales with the purpose to cover the B2B-specific purchase and decision criteria. They, thus, differ from the scales that researchers use for B2C brands (e.g., Fischer, Völckner, and Sattler 2010). For the validation, we apply a confirmatory factor analysis. Results confirm our concepts. Indeed, the brand function shapes the relevance of B2B brands. As expected, this function is reflected by the two subfunctions risk reduction and imagery benefits. We, hence, extend previous conceptualizations of B2B brands as, for example, the framework of Brown et al. (2011). While these frameworks do not include imagery benefits of brands, we confirm that imagery benefits are a valid subfunction of brands.

We gain additional insights from the measurement of B2B brand relevance. While B2B brand relevance strongly differs between Germany and the US, differences between product categories are somewhat smaller. Additional variance exists across individual managers. We, hence, conduct a separate driver analysis for B2B brand relevance in Germany and the US. Here, we particularly test how individual manager and organizational characteristics moderate the influence of the brand functions.

For Germany and the US, we can show that both subfunctions, risk reduction and imagery benefits, drive B2B brand relevance. Thus, we are the first to actually show that imagery benefits drive B2B brand relevance. On average, however, risk reduction has a stronger impact than imagery benefits. Although image components of brands are important for managers, brands' potential to reduce risks seems much more important for them.

Concerning the moderators, we find remarkable differences between Germany and the US. For German managers, risk reduction aspects have lower importance and imagery bene-

fits have higher importance for brand relevance, when they take responsibility for their actions alone compared to a situation when they are one of several responsible group members. The opposite pattern holds for the US.

Our study also offers various insights for B2B managers. They can use our scale to measure brand relevance and to align their branding strategies accordingly. Moreover, our concept and driver analysis shows that brand functions on B2B markets address information efficiency concerning risk reduction and imagery benefits. Although risk reduction seems to be more important, managers should not neglect the imagery components of B2B brands. They may be crucial for a purchase decision. As the impact of both brand functions varies strongly with manager characteristics, these characteristics may lead to a comparably higher impact of imagery components. Managers of the brand-owning company can use our detailed driver analysis to gain insights on which brand function to strengthen for which manager of a potential purchasing company.

3.4 Project 4: Culture and Survey Research: A Review on How Response Styles Differ
across Cultures

Today's world is characterized through globalization and migration. Hence, it becomes increasingly common that people from various cultures live at one geographical place. For example, in 2015, more than 20% of the German population have had a migration background (Statistisches Bundesamt, 2016). Other countries show comparable figures and numbers are still increasing. Simultaneously, the number and impact of cross-cultural studies increases.

Both developments may especially bias the results of questionnaires. Here, respondents tend to apply specific response styles in their answer behavior. For example, they tend to choose the midpoint of the answer categories regardless of the content of the item (for an overview on response styles see, e.g., Baumgartner and Steenkamp 2001). Hereby, individual respondent characteristics promote certain response styles. In particular, culture is one factor

that shapes whether a respondent shows a certain response style. For example, respondents from Hispanic cultures seem to choose more extreme answer categories than respondents from Anglo cultures (e.g., Hui and Triandis 1989). Importantly, these response styles can contaminate measurements and their interrelations (Baumgartner and Steenkamp 2001; Greenleaf 1992). If researchers and managers do not consider the effect of culture on respondents' response styles, they may easily draw wrong conclusions from survey data (Schwarz 2003). This problem is especially vital because migration is likely to increase.

In turn, such wrong conclusions may endanger companies businesses. Companies use survey data to capture various latent constructs, such as brand experience (e.g., Brakus, Schmitt, and Zarantonello 2009) and price consciousness (e.g., Lichtenstein, Ridgway, and Netemeyer 1993). As shown by the stimulus-organism-response model, these parameters explain the link between companies' actions and consumer reactions. They carry those insights that companies use for their strategic planning. Therefore, if companies do not consider response styles according to culture, they may draw wrong conclusions. These wrong conclusions may lead to substantially lower profits.

The first objective of this study is, hence, to develop a framework from literature that depicts how culture influences response styles. I develop my framework based on the framework of Tourangeau, Rips and Rasinski (2009), which splits up the response process into comprehension, retrieval, judgment, and response. Culture may influence all of these steps (e.g., Johnson et al. 1997) so that respondents with different cultural backgrounds show different levels of response styles. Thereby, culture includes beliefs, values, and norms shared with a person's initial sociocultural environment and those beliefs, values, and norms learned by confrontation with other cultures. Item characteristics moderate this process. They may reduce or increase the extent to which cultures differ in their tendency to show a specific response style.

The second objective of the project is to derive generalizations for response styles due to culture. Thus, I assess existing literature on how different cultures shape response styles. I consider seven common response styles. Reviewed articles compare effects of cultures between countries, between immigrants and indigenes, along cultural dimensions, and finally moderations of these effects by item design. I assess these articles concerning convergent and divergent findings and derive generalizations on response styles. For example, in comparison to individualistic cultures, collectivistic cultures seem to promote an acquiescence response style where respondents tend to agree on items regardless of their content. I sum up generalizations in several tables (See project 4 table 3 (p. 161), table 4 (p. 168), and table 5 (p. 173)). Companies and researchers can use these generalizations as intuition on whether the cultural composition of their samples may foster differences in response styles that contaminate their results. This notion is especially important for cross-cultural research. Moreover, managers and researchers may use findings for the moderation effects of item characteristics to adopt initial item design. For example, in comparison to a standard Likert scale format, an interrogative scale format seems to reduce differences in acquiescence response styles between Asian and Anglo cultures.

For certain groups, however, findings are scarce and even contradictive. For example, studies show that certain scale formats may reduce cultural differences in response styles. However, these studies are rare and not generalizable in terms of the cultural composition of their samples. Thus, the third objective of this project is to point out promising avenues for further research. In detail, I derive seven research gaps. The first three research gaps address the interplay of culture and item design. Here, researchers should especially test how scale formats may reduce differences in response styles due to culture or acculturation. Research gap four to six take into account a new or more detailed distinction of culture. In particular, future studies should consider the effects of business culture on response styles in manager

surveys. Finally, research gap seven demands for the replication of results with marketing scales.

Overall, results of this study reinforce the view that culture may shape response styles and, thus, contaminate results. Managers and researchers, thus, need to carefully assess their samples. This notion is not only true if managers and researchers conduct cross-cultural studies, but concerns all studies. For example, even if they draw their sample in one country, samples may consist of immigrants and indigenes that strongly differ in cultural backgrounds. Managers and researchers, hence, have to carefully check their sample compositions and check for according contamination of results. Otherwise, they may draw wrong conclusions which may even negatively affect a company's profit.

REFERENCES SYNOPSIS

Bates, John M. and André Gabor (1986). "Price Perception in Creeping Inflation: Report on an Enquiry," *Journal of Economic Psychology*, 7 (3), 291-314.

Baumgartner, Hans and Jan-Benedikt E. M. Steenkamp (2001). "Response Styles in Marketing Research: A Cross-national Investigation," *Journal of Marketing Research*, 38 (2) 143-156.

Bijmolt, Tammo H., Harald J. van Heerde, and Rik G. M. Peters (2005). "New Empirical Generalizations on the Determinants of Price Elasticity," *Journal of Marketing Research*, 42 (2), 141-156.

Brakus, J. Joško, Bernd H. Schmitt, and Lia Zarantonello (2009). "Brand Experience: What Is it? How Is it measured? Does it Affect Loyalty?," *Journal of Marketing*, 73 (3), 52-68.

Brown, Brian P., Alex R. Zablah, Danny N. Bellenger, and Wesley J. Johnson (2011). "When do B2B Brands Influence the Decision Making of Organizational Buyers? An Examination of the Relationship between Purchase Risk and Brand Sensitivity," *International Journal of Research in Marketing*, 28 (3), 194-204.

Brown, Tom J. and Peter A. Dacin (1997). "The Company and the Product: Corporate Associations and Consumer Product Responses," *Journal of Marketing*, 61 (1), 68-84.

Fischer, Marc, Wolfgang Giehl, and Tjark Freundt (2011). "Managing Global Brand Investments at DHL," *Interfaces*, 41 (1), 35-50.

Fischer, Marc, Franziska Völckner, and Henrik Sattler (2010). "How Important Are Brands? A Cross-Category, Cross-Country Study," *Journal of Marketing Research*, 47 (5), 823-839.

Greenleaf, Eric H. (1992). "Measuring Extreme Response Styles," *Public Opinion Quarterly*, 56 (3), 328-351.

Homburg, Christian, Martin Klarmann, and Jens Schmitt (2010). "Brand Awareness in Business Markets: When Is it Related to Firm Performance?," *International Journal of Research in Marketing*, 27 (3), 201-212.

Homburg, Christian, Wayne D. Hoyer, and Nicole Koschate (2005). "Customers' Reaction to price Increases: Do Customer Satisfaction and Perceived Motive Fairness Matter?," *Journal of the Academy of Marketing Science*, 33 (1), 36-49.

Hui, Harry and Harry C. Triandis (1989). "Effects of Culture and Response Format on Extreme Response Style," *Journal of Cross-Cultural Psychology*, 20 (3), 296-309.

Interbrand (2016). *Best Global Brands 2015* [available at www.interbrand.com; accessed at November 4, 2016].

Johnson, Timothy, Diane O'Rourke, Noel Chavez, Seymour Sudman, Richard Warnecke, Loretta Lacey, and John Horm (1997). "Social cognition and responses to survey questions among culturally diverse populations," in L. Lyberg, P. Biemer, M. Collins, E. de

Leeuuw, C. Nippo, N. Schwarz, D. Trewin (Eds.), *Survey Measurement and Process Quality*, New York: Wiley, 87-113.

Katona, George (1974). "Psychology and Consumer Economics," *Journal of Consumer Research*, 1 (1), 1-8.

Keller, Kevin L. (1993). "Conceptualizing, Measuring, and Managing Customer-Based Brand Equity," *Journal of Marketing*, 57 (1), 1-22.

Lichtenstein, Donald R., Nancy M. Ridgway, and Richard G. Netemeyer (1993). "Price Perceptions and Consumer Shopping Behavior: A Field Study," *Journal of Marketing Research*, 30 (2), 234-245.

Ownbey, Shiretta F. and Patricia E. Horridge (1997). "Acculturation Levels and Shopping Orientations of Asian-American Consumers," *Psychology and Marketing*, 14 (1), 1-18.

Peñaloza, Lisa N. (1994). "Atravesando Fronteras/Border Crossings: A Critical Ethnographic Exploration of the Consumer Acculturation of Mexican Immigrants," *Journal of Consumer Research*, 21 (1), 32-54.

Schwarz, Norbert (2003). "Self-Reports in Consumer Research: The Challenge of Comparing Cohorts and Cultures," *Journal of Consumer Research*, 29 (4), 588-594.

Sethuraman, Raj, Gerard J. Tellis, and Richard A. Briesch (2011). "How Well Does Advertising Work? Generalizations from Meta-analysis of Brand Advertising Elasticities," *Journal of Marketing Research*, 48 (3), 457-471.

Statistisches Bundesamt (2016). *Migration und Integration* [available at http://www.destatis.de; accessed November 4, 2016].

Tourangeau, Roger T., Lance J. Rips, and Kenneth Rasinski (2009). *The Psychology of Survey Research*, Cambridge: Cambridge University Press.

Wong, Nancy, Aric Rindfleisch, and James E. Burroughs (2003). "Do Reverse-worded Items Confound Measures in Cross-cultural Consumer Research? The Case of the Material Values Scale," *Journal of Consumer Research*, 30 (1), 72-91.

Woodworth, Robert S. (1929). *Psychology*, New York: Holt.

Zablah, Alex R., Brian P. Brown, and Naveen Donthu (2010). "The Relative Importance of Brands on Modified Rebuy Purchase Situations," *International Journal of Research in Marketing*, 27 (3), 248-260.

PROJECT 1: CONSUMERS' PERCEPTION OF PRICE INFLATION

Authors: Eric Michael Lennartz and Marc Fischer

CONTRIBUTION STATEMENT

Although perceived inflation has strong impacts on consumer behavior (e.g., Katona 1974), almost no study from consumer research considers this phenomenon. Yet, no generalizable method to measure perceived inflation exists. Thus, we develop and validate a general approach to measure perceived inflation on a global psychological scale. We base this approach on actual perceptions only. Thereby, we contribute to research from a methodological perspective. Researchers and managers can use our approach to track perceived inflation and to gain insights on corresponding consumer reactions. Moreover, research lacks knowledge on how consumers form inflation perceptions and what are potential drivers of this process. We conceptualize this formation and analyze its drivers. Here, we substantially contribute to research by showing that the perception of inflation consists of two steps: consumers first perceive product price trends that deviate from official figures and, second, weight them based on a subjective market basket. We separate both steps in our driver analysis. Thus, we can show which factors drive the fact that perceived inflation deviates from official figures and which of the two steps each factor affects.

ABSTRACT

Statistical offices put strenuous efforts into the measurement of price inflation. On the contrary, consumers' perception of inflation strongly deviates from official measures and may shape price fairness, price sensitivity, and shifts in consumption. Therefore, we develop and validate a psychological rating scale to measure perceived inflation. We apply this scale to 13 product groups and 53 product classes. We find that consumers tend to overestimate inflation, but that the individual perception largely depends on consumer and category characteristics. Moreover, we conceptualize the formation process of perceived inflation and analyze its drivers. Here, we separate two processes, consumers' perception of individual price changes and their weighting according to consumers' perceived market baskets. We show which factors drive the formation of perceived inflation and which process they affect. For example, purchase frequency has opposite effects on both processes. As perceived inflation may shape consumer behavior, we offer a tool that managers and researchers can use to track perceived inflation. Moreover, our driver analysis provides insights on which price changes for individual products are most likely to influence overall inflation perceptions.

Keywords: Inflation; Prices; Price Perception; Scale Development; Price Knowledge.

1   Introduction

Statistical offices put strenuous efforts into the measurement of price inflation. The Bureau of Labor Statistics collects data from about 23,000 establishments every month (Bureau of Labor Statistics 2016). Likewise, consumers' perception of inflation strongly deviates from these measures (e.g., Bates and Gabor 1986). When Germany introduced the Euro as its new currency in 2002, the official inflation rate was modest at 1.4%. Nonetheless, Germans called the new currency "Teuro". They combined the German word for expensive "teuer" with the name of the new currency. Thus, as official inflation only partly meets consumers' perception, working with official inflation rates may easily lead to false and biased results for consumer research.

Meanwhile, perceived inflation has strong impact on research and real-world decisions. For low reputation brands, perceived inflation might lower price fairness evaluations (Xia, Monroe, and Cox 2004). Moreover, high inflation perceptions may increase price sensitivity (Bijmolt, van Heerde, and Pieters 2005). Finally, unexpected perceived inflation reduces or shifts consumption in a category (Katona 1974). Tracking perceived inflation, hence, enables economists and managers to forecast and to react to market developments and structures.

Our study, thus, has important methodological contribution. All existing methods to measure perceived inflation serve specific purposes. Thus, they come with assumptions that limit their scope. Therefore, our research first addresses how to measure perceived inflation on a global psychological scale. Based on limitations of current approaches, we develop a psychological scale that researchers, economist, and managers can use to track perceived inflation. It offers a more precise picture of consumer behavior in comparison to official inflation figures. We validate and apply it to 13 product groups and 53 product classes, including services, durables, and nondurables. However, scale values are not comparable to actual percentage figures for official inflation. Thus secondly, we estimate a function to map the per-

ceived inflation scale for a category on a percentage scale that is comparable to the official inflation rate.

Statistical offices weight product inflation rates by a market basket, which includes average consumer expenditures, to construct the overall rate of inflation. Consumers' perceived product price changes and market baskets differ from official figures. They subjectively weight the importance of products for overall inflation (e.g., Bruine de Bruin, van der Klaauw, and Topa 2011) and perceive prices inaccurately (e.g., Vanhuele and Drèze 2002). These deviations translate to deviations on higher aggregation levels. Research has neither analyzed this process in detail, nor its drivers. Therefore thirdly, we analyze the drivers of perceived product price changes and their weights for category perceived inflation. We measure perceived inflation for product classes and estimate the composition of consumers' "perceived market baskets". Both differ from official figures. Hence, we conduct detailed driver analyses for both.

Hereby, we offer substantive contribution to the understanding and conceptualization of perceived inflation. We can show that consumers largely over- and underestimate inflation. This deviation differs between the consumers. We detail the knowledge on the formation of perceived inflation and show how it arises from individual perceptions of product price changes and their aggregation weights. By separating the perception of product price trends and weights in the market basket, we can give a more precise picture than previous studies. In fact, we can show in which process particular drivers influence the perception of inflation.

We structure our study as follows. We first explain the concept of perceived inflation and illustrate our model and methodology. Then, we refer to the measurement of perceived inflation and the driver analyses. Finally, we discuss results, derive implications, and give a conclusion.

2 The Concept of Perceived Price Inflation

In line with existent literature (e.g., Jonung 1981), we define perceived inflation as an aggregate value distinct for each person that considers her/his overall evaluation of all single price changes within a time period. This definition is equivalent to the official inflation measurement except that all evaluation processes have underlying subjective perceptions.

Related literature comes from economics, consumer behavior, and behavioral economics. In economics, research deals with the construction and requirements of inflation indexes (e.g., Fisher 1922). These indexes usefully aid economic decisions. However, they assume that consumers are rational and, hence, cannot explain how individuals perceive inflation. In contrast, consumer researchers address how individuals perceive, evaluate, and learn prices. Nevertheless, among them only Katona (1974) and Bolton, Warlop, and Alba (2003) touch on perceived inflation.

Behavioral economists combine both views. They address the psychology behind perceived inflation (e.g., Ranyard et al. 2008). Table 1 gives an overview of these studies. In particular, studies use four approaches to measure perceived inflation. Brachinger (2008) mathematically corrects the inflation index based on external assumptions for consumers' perceived market baskets. In a second approach, consumers state percentage estimates of perceived inflation (e.g., Jonung 1981). The third approach uses price change tendencies, e.g., by the EU Consumer Survey (2016), and transforms them into percentage values based on the assumption that consumers track the medium trend of inflation correctly (e.g., Dräger, Menz, and Fritsche 2014). For the fourth approach, finally, consumers state the price of a specific product at the beginning and the end of a measurement period, for example, today and three months ago (e.g., Bates and Gabor 1986). Based on these values, researchers can construct perceived inflation rates for single products.

Table 1: Overview on Studies Measuring Perceived Inflation

Study	Jonung (1981)	Bates and Gabor (1986)	Jonung (1986)	Jonung and Laidler (1988)	Antonides (2008)	Brachinger (2008)	Dräger, Mentz, and Fritsche (2014)	Our Study
	American Economic Review	Journal of Economic Psychology	Journal of Economic Psychology	American Economic Review	Journal of Economic Psychology	Journal of Economic Psychology	Applied Economics	
Measurement								
Approach	Percentage statement	Percentage statement / Price statements	Price change tendency / Percentage statement	Percentage statement	Price change tendency	Statistical Correction	Price change tendency with transformation	Likert Scale with transformation
Number of items	single item	single item	single item	single item	single item	-	single item	multi item
Construct validation	-	-	-	-	-	-	-	✓
Separation of measurement and transformation into percentage values	-	-	-	-	-	-	✓	✓
Level of assumptions	low	low	low	low	low	high	medium	low
Substantive drivers								
Consumer demographics	✓	-	-	-	-	-	-	✓
Consumer psychographics	-	-	-	-	-	-	-	✓
Category characteristics	-	-	-	-	-	-	-	✓

These indexes fit to specific purposes and, hence, use tailored assumptions. Our approach aims at a more general understanding. Drawing from the second to the fourth approach, we measure consumers' actual inflation perceptions. Thus, we do not need assumptions for the perceived market basket. However, contrary to these approaches, we use a multi-item scale and validate it according to classical test theory. This procedure enables us to get a more precise picture of perceived inflation (deVellis 2012). In line with the third approach, we also consider that consumers have limited abilities to deal with percentage values (Chen and Rao 2007) and separate the measurement of perceived inflation from its transformation into percentage values. But, as consumers struggle with tracking individual prices correctly (e.g., Vanhuele and Drèze 2002), we do not assume that they track the medium inflation rate correctly, but build our transformation on their actual perceptions.

Further, studies show that perceived inflation results from two kinds of deviations, a deviation of perceived market basket and a deviation of perceived product price trends from official figures (e.g., Antonides 2008; Bates and Gabor 1986). Although Jonung (1981) also regresses overall perceived inflation on consumer demographics, to the best of our knowledge, no study analyzes what drives perceived product price trends and market baskets. We overcome this lack of research and conduct a driver analysis for both. In addition to Jonung (1981), we also include consumer psychographics and category characteristics as drivers.

3   A Model to Measure and Explain Perceived Inflation

3.1 Consumers' Perception of Price Inflation

Figure 1 exemplarily illustrates how statistical offices measure inflation for the product group Recreational and Cultural Services. They measure price changes for individual products and a market basket that includes consumers' average expenditures for a product in relation to

their overall expenditures. Statistical offices use this basket to weight the individual product price changes and aggregate them to inflation rates. In such a way, they determine inflation rates for all aggregation levels of the Classification of Individual Consumption according to Purpose (COICOP; for the overall scheme see United Nations Statistics Division (2016)), such as product classes and groups. To perceive price trends as given by official figures, consumers would have to track both product price trends, that is past and current product prices, and weights in the market basket correctly. In contrast, they perceive product prices inaccurately (e.g., Vanhuele and Drèze 2002) and use heuristics to aggregate prices (e.g., Desai and Talukdar 2003).

Figure 1: Inflation Measurement for the Product Group Recreational and Cultural Services
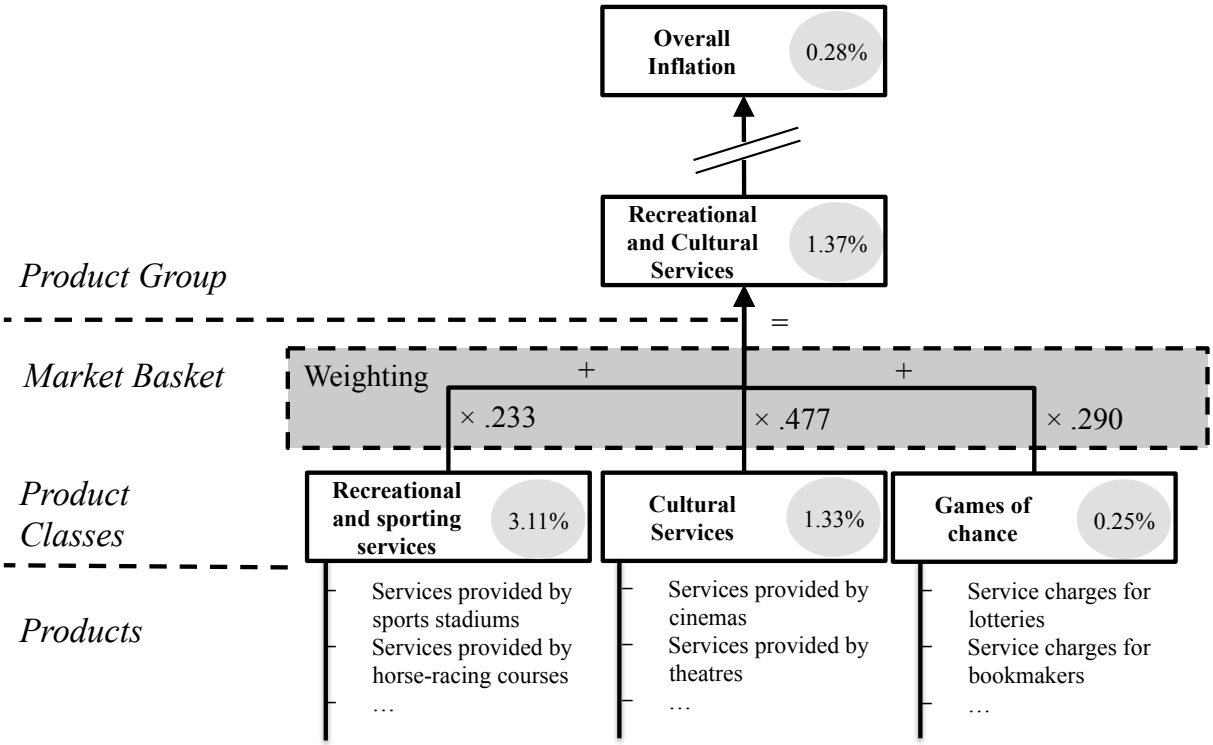
According to the multi-store model of memory (Atkinson and Shiffrin 1968), price information enters a consumer's memory as sensory inputs. Hence, price exposure is necessary to learn prices (Monroe and Lee 1999) and, thus, price trends. A misperception of a product price trend may, thus, result from a lack in price exposure. For example, the German Federal

Statistical Office (2016) measures more than 300,000 prices a month. Consumers can only come across a small percentage of these prices.

Price exposure depends on shopping patterns (e.g., Estelami and Lehmann 2001), engagement into price search (e.g., Vanhuele and Drèze 2002), and external stimuli (e.g., Estelami and de Maeyer 2004). In addition, prices for a single product can change unlike the overall price trend in a product class. Exclusive exposure to the prices of selected products may enhance incomplete exposure. If so, although consumers perceive the price change of these products more accurately, larger biases may emerge for overall price trends (e.g., Bruine de Bruin, van der Klaauw, and Topa 2011).

Likewise, much information enters a consumer's memory simultaneously. Price exposure is, thus, not sufficient for price knowledge. If consumers do not pay sufficient attention to new price information, it decays and does not enter the working memory where consumers encode information into the long-term memory (Attkinson and Shiffrin 1968). However, only if consumers store information in the long-term memory, they may retrieve it when evaluating inflation.

In the working memory, consumers rehearse information. Overall, rehearsal underlies limited cognitive resources of the working memory. For price information, it occurs as conscious or nonconscious processing and involves many comparative judgments (e.g., Monroe and Lee 1999). Consumers retrieve information from long-term memory for these judgments. Based on these judgments, they construct new cues in the long-term memory, replace existent cues, form associations between existent and new information, or the new information decays. Consumers perceive product price trends when they combine new with past price cues of a product.

Thus, deviations of official and perceived product price trends may also arise from the fact that consumers may not process information thoroughly. Consumers may put in only low

efforts or fail to retrieve and encode information correctly. In addition, cognitive load may be too high to process all information at one point in time (Vanhuele and Drèze 2002).

Moreover, during rehearsal, consumers may combine several product price trends to aggregated price trends, for example, for product groups or on an overall level. For this aggregation, they weight information according to their individual perceived relative importance (Bruine de Bruin, van der Klaauw, and Topa 2011). This importance may deviate from the official weights in the market basket. Here, consumers form their perceived market baskets.

Due to limited cognitive resources, consumers use the representativeness and availability heuristic when they evaluate information. According to the representativeness heuristic (Kahneman and Tversky 1972), consumers perceive an event as more probable when it is perceived as representative for a class of events. The availability heuristic (Tversky and Kahneman 1973) connects a higher probability to an event that is more easily accessible. Therefore, similar to the formation of store price images, we expect that the relative importance of a price change and, thus, its weight for the aggregation process increases in its (1) availability, (2) salience, and (3) comparability (Lourenço, Gijsbrechts, and Paap 2015).

3.2 Basic Idea of our Model

Following our argumentation above, we expect that perceived inflation arises from two steps: (1) Consumers perceive price changes for individual products and (2) aggregate them according to their perceived relative importance to aggregated price trends, e.g., for a product group. The weights for the aggregation form a consumers' perceived market basket. This aggregation process is supposedly often implicit.

Due to limited cognitive capacities, the aggregation process in the consumers' memory has to involve various intermediate aggregation levels. Equivalent to the official inflation measurement, we assume that consumers aggregate individual price trends similar to the structure of the COICOP as given in figure 1. For example, we expect them to first aggregate

individual price trends to price trends for product classes and then conduct further aggregations. Within this scheme, we focus on the aggregation step between the product class and group. We choose this step because product classes are generic and graspable for consumers.

Figure 2 depicts our conceptualization including the underlying processes and driver categories. According to Estelami and Lehmann (2001), drivers of price recall accuracy come from the product category, consumer psychographics, the economic environment, consumer demographics, and task characteristics. Similarly, we expect variables from the product category, consumer psychographics, and the economic environment to drive both, the deviation of perceived inflation from official inflation in a product class and the weights in the perceived market basket. We divide drivers from the product category along initial product category characteristics and price development. We include consumer demographics and task characteristics as control because they only affect the deviation of perceived and official inflation in a product class.

Figure 2: Conceptual Model for the Drivers of Perceived Inflation

29
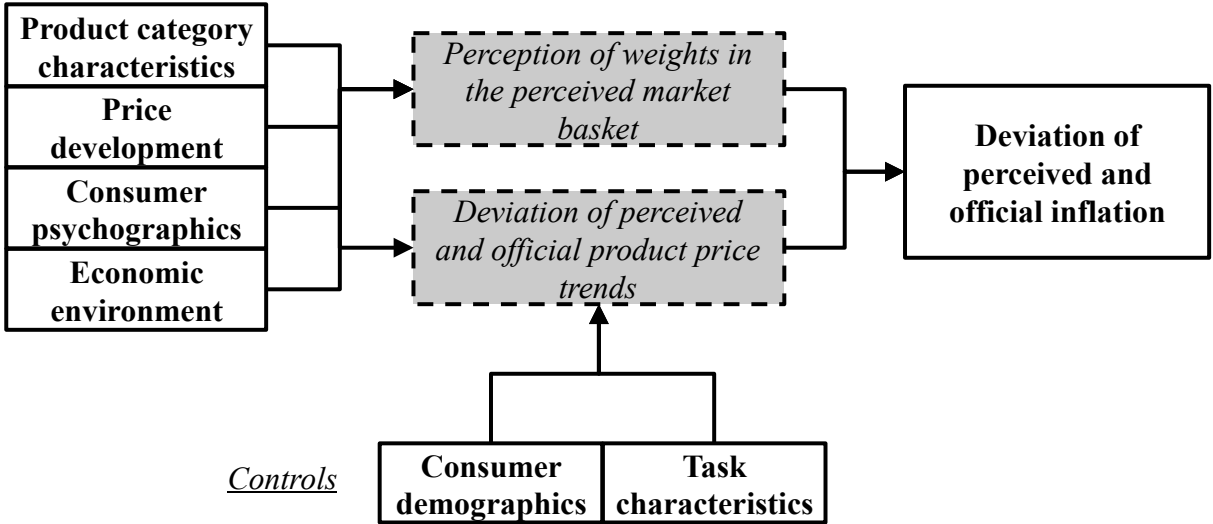
3.3 Drivers of the Deviation of Perceived and Official Price Changes for Product Classes

3.3.1 Product Characteristics

A higher purchase frequency in a product class leads to a higher level of price exposure (Estelami and de Maeyer 2004). Similarly, recent purchases in a product class enhance the probability of exposure to recently adapted prices. Hence, we expect that biases in price exposure over time decrease. In addition, nondurables come with lower prices and involvement than durables and services (Estelami and de Maeyer 2004). For durables and services, consumers may engage more into price search so that the level of price exposure may increase. However, as consumers buy durables infrequently, this engagement may occur only at one point in time. Consumers may dismiss to track price movements over time. Finally, a broader category clutter induces variety in price changes across a product class and, hence, higher cognitive load. Thus, deviation of perceived and official price changes is likely to be lower for product classes with a high *purchase frequency*, *purchase recency*, and *nondurables* and *services* in comparison to durables and higher for a broader *category clutter*.

3.3.2 Price Development

For a single price, a large price range leads to confusion and high cognitive load (Vanhuele and Drèze 2002). Moreover, for a high perceived-price promotion frequency, consumers often experience price changes and get a feeling for the correct price change size (Vanhuele and Drèze 2002). Biases due to incorrect retrieval decrease. We assume that a higher *price change range* increases, while a higher *perceived price promotion frequency* reduces the deviation of perceived and official inflation in a product class.

3.3.3 Consumer Psychographics

More active price search enhances a consumer's price exposure. Low search costs online, using prices as quality signal, and price consciousness enforce this price search (Lichtenstein, Ridgway, and Netemeyer 1993; Lynch and Ariely 2000). However, although price conscious consumers may have a higher level of price exposure, they may only come across prices of low-priced products that reveal atypical price changes. Contrary, brand loyalty and expertise in a product class decrease price search because they reduce the relative importance of the price for a purchase. Consumers' *affinity for online shopping* and the usage of the *price-quality relationship* lowers the deviation of perceived and official inflation in a product class, while *price consciousness*, *brand loyalty*, and *customer expertise* increase it.

3.3.4 Economic Environment

Consumers gather knowledge on the economic environment from media and their personal environment. Both strengthen selected information, which is called social amplification (Ranyard et al. 2008). In this context at least, media reports more about increasing than decreasing prices (Soroka 2006). Hence, we assume that *social amplification* of price changes increases the deviation of official and perceived inflation in a product class.

3.4 Drivers of the Weights in the Perceived Market Basket

3.4.1 Product Characteristics

Purchase frequency and recency may increase the availability and salience of price changes (Ranyard et al. 2008). Similarly, a broader category clutter represents more prominent information and enhances the salience of product price changes (Lourenço, Gijsbrechts, and Paap 2015). In addition, product classes for which consumers spend more at an individual shopping trip attract focal attention during price search and become more salient (Desai and

31

Talukdar 2003). For *purchase frequency*, *purchase recency*, *category clutter*, and *monetary impact* we expect a positive effect on the weight of a product class.

3.4.2 Price Development

A higher range of price changes may reduce comparability of price changes in a product class. In contrast, more frequent price promotions may stimulate higher attention and a more accurate processing of price trends (Desai and Talukdar 2003; Vanhuele and Drèze 2002). Price changes become more available and salient. Larger price changes may also be more salient because they have stronger impact on consumer purchase evaluations (Ranyard et al. 2008). According to loss aversion, price increases loom larger than price decreases and are, thus, more salient. Both effects may multiply and reveal a positive interaction. The official inflation rate equals this interaction. Overall, a larger *price change range* in a product class is likely to decrease the weight in the perceived market basket, while a higher *perceived price promotion frequency*, a positive *direction of a price change*, a larger *size of a price change*, and a higher *official inflation rate* in a product class are likely to increase it.

3.4.3 Consumer Psychographics

If consumers mainly shop for a product class online, their probability to recognize price information and, thus, the salience of price changes is higher due to lower search costs. Similarly, product classes that refer to a higher price consciousness experience increased attention so that price changes become more salient (Lichtenstein, Ridgway, and Netemeyer 1993). In addition, if consumers perceive expertise in a field, we expect price trends to be more salient and available for them. We assume that higher *affinity for online shopping*, *price consciousness*, and *customer expertise* lead to a higher weight of a product class in the perceived market basket.

3.4.4 Economic Environment

Social amplification enhances the availability of price changes (Ranyard et al. 2008) and is more prevalent for price increases than decreases (Soroka 2006). We expect *social amplification* to increase the weight of a product class in the perceived market basket and to show a positive interaction with the *direction of price chang*e.

4 Methodology

Figure 3 illustrates our methodology. We follow a four-step approach with two surveys (survey 1 and survey 2). First, we develop and validate a global rating scale for perceived inflation. In survey 1, we apply this scale to 13 product groups. As we conduct our survey in Germany, we choose product groups from the official market basket of the German consumer price index, which follows the COICOP (for an overview of the selected groups and classes see appendix A). We restrict our choice to groups that contain unambiguous and familiar product classes and for which the German Federal Statistical Office assesses three to six underlying product classes. Whereas, for groups with fewer classes, the weights in the market basket are straightforward, our groups promise a sufficient complexity and, thus, insights into consumers' perceived market baskets. At the same point, our group choice avoids biases from lengthy questionnaires as for groups with more classes. In addition, it promises generalizable findings because these groups represent 35.17% of consumers' purchase volume (German Federal Statistical Office 2016) and cover nondurables, durables, and services.

Second, we use survey 2 to estimate a function that transforms scale values for perceived inflation into percentage values. In a quasi-experiment, consumers rate price changes on our perceived inflation scale. These ratings enable us to estimate a function that translates rating values into percentage values. Moreover, this approach considers consumers' abilities by avoiding that survey participants have to deal with percentage numbers.

Figure 3: Overview of our Research Approach

Research Process	Purpose	Survey	Model
1. Global Inflation Scale Scale Development	• Measurement of Perceived Inflation • Validation	• Survey 1	
2. Mapping Function Quasi-Experiment	• Compare Perceived Inflation with Official Inflation	• Survey 2	• Model for measuring category perceived inflation • equations (1), (2), and (3)
3. Driver Analysis for the Perceived Price Trends for Product Classes Driver Analysis	• Understanding the Deviation of Official and Perceived Product Price Trends	• Survey 1	• Model of drivers of deviations of perceived and official price changes in product classes • equation (8) – (9)
4. Driver Analysis for the Weights in the Perceived Market Basket Driver Analysis	• Understanding the Relative Importance of Product Classes	• Survey 1	• Model of drivers of weights in the perceived market basket • equation (10) – (11)

We analyze, third, what drives the deviation of perceived from official price changes in a product class and, fourth, what drives weights in the perceived market basket. Therefore in survey 1, respondents rate perceived inflation and several potential drivers for the product classes. From these ratings, we determine the deviation of perceived inflation and official inflation. In addition, we determine the weights in the perceived market basket from a regression of perceived inflation for a product group on perceived inflation in the corresponding product classes. Finally, we regress both, deviations and weights, on their potential drivers.

5 Measurement of Perceived Inflation

5.1 Development of a Rating Scale

5.1.1 Item Generation and Scale Purification

First, we generated an item battery based on literature (e.g., Monroe 2003). We use an 11-point Likert scale with answer categories that range from -5 (do not agree at all) to 5 (totally agree). Negative labels of answer categories mark deflation, positive labels mark inflation. In pre-studies, we discussed and evaluated our items with colleagues from market research and potential respondents to ensure content and face validity. Afterwards, we pretested our items for a subsample of product groups with two convenience samples of 118 and 224 respondents. Considering scale reliability and construct validity, we reduced the scale from initially 37 to three items. Appendix B provides further details on this process. Table 2 shows our resulting three-item scale. It covers three processes to change prices: changing (nominal) prices, quantity, and quality of a product or service (Monroe 2003).

Table 2: Translations of Measures within the Questionnaire[1]

Perceived inflation[2]
    Prices have risen strongly within the last 12 months.
    The same consumption of products and services costs significantly more today than 12 months ago.
    Today, I pay substantially more for the same quality than 12 months ago.

1) original scale items were German; 2) 11-point Likert scale (-5: I do not agree at all, …, 5: I totally agree).

5.1.2 Data and Questionnaire

Tables C1 and C2 in appendix C provide all scales of survey 1, their sources, and the original items in German. Each respondent rates exactly one product group. First, s/he rates our perceived inflation scale for the given product group and corresponding validity measures. Further, s/he rates perceived inflation for the product classes of that product group.

We describe product groups and classes by a list of products to ensure an equal understanding. Finally, we measure product class drivers via single item scales. Here, we choose single items to avoid biases due to mental fatigue and boredom (deVellis 2012).

We collected data for survey 1 via an online-survey between April 14th and 23rd, 2015 with respondents from an existing online panel. All respondents live in Germany, exceed the age of 16 years, and have purchased in a corresponding product group within the last 12 months. To ensure data quality, we screened out respondents that failed a bogus item. Our sample includes 1,439 respondents. 507 respondents are female (42.89%; percentage values are based on those respondents that reported a value for a demographic characteristic), 675 are male (57.11%). Mean age is 44.32 (SD = 14.54) and mean net income is 2133.62 € (SD = 1268.58). 367 respondents (29.43%) are university graduates and 113 (9.10%) are students. Distributions of demographic characteristics in our sample slightly differ to the overall German population. This trend holds when we consider demographics of the samples for the individual product groups. Hence, we weight observations according to their representativeness when we determine mean values of perceived inflation and average weights in the perceived market basket. In addition, we control for demographics when we estimate our mapping function.

5.1.3 Reliability

Internal consistency is excellent. For all aggregation levels, coefficient alpha is .949 on the overall level and exceeds .900 for all product groups and classes. We conduct a confirmatory factor analysis. Average variances extracted from our items lie between .768 and .979 for the overall, product group and product class level. Factor loadings are all highly significant (t-values >10). Composite reliabilities range from .908 to .992.

5.1.4 Convergent Validity

National consumer surveys, as the EU Consumer Survey (2016), measure perceived inflation by a price change tendency and a percentage statement (for detailed measures, see appendix C). We include both measures in survey 1 to test convergent validity. We also include the official inflation rate for the product classes from the German Federal Statistical Office as an external measurement to assess convergent validity. Table 3 reports the correlation between perceived inflation and the convergent validity measures. We correct them by their measurement error, if available (deVellis 2012)[1]. As all of these correlations are positive and significant, they support convergent validity.

Table 3: Convergent and Nomological Validity Measures

Measure	#Items	N	Coefficient Alpha	Correlation with Perceived Inflation[1]
Perceived Inflation per Product Group	3	1,439	.949	-
Convergent Validity (internal)				
Perceived Price Change Tendency[2]	1	1,408	-	.617***[4]
Perceived Inflation by Percentage Statement	1	1,394	-	.588***[5]
Convergent Validity (external)				
Official Inflation Statistics (Product Classes)	-	52	-	.319**[6]
Nomological Validity (internal)				
Expected Inflation	3	1,345	.963	.743***
Consumption Intention[3]	3	1,329	.863	-.296***
Saving Intention	4	1,333	.687	.453***
Actual Postponement of Purchases	1	1,333	-	.328***
Intended Postponement of Purchases	1	1,333	-	.440***
Price Fairness	6	1,333	.908	-.625***

1) correlations adjusted by reliability (see footnote 1); 2) non-parametric coding (-1: fallen, 0= stayed about the same, 1=slightly/moderately/strongly risen); 3) values are reverse scored to the original scale; 4) value represents Spearman-Rho as the variable perceived price change tendency is ordinal; 5) price increases over 50% were excluded; 6) values are based on product group/class averages weighted according the representativeness of individual observations for overall population of Germany; *$p < .1$, **$p < .05$, ***$p < .01$ (two-sided test).

---

[1] Only the reliable parts of indicators can correlate (deVellis 2012). Thus, if r_{aa} and r_{bb} are the reliabilities of two constructs A and B, their correlation is restricted to $\pm\sqrt{r_{aa}r_{bb}}$. We count for this attenuation and correct scores by multiplying them with $1/\sqrt{r_{aa}r_{bb}}$.

5.1.5 Nomological Validity

Theory postulates inflation expectations, saving intentions, and intended and actual postponement of purchases to be positively related to perceived inflation (Katona 1974; Ranyard et al. 2008). Price fairness and consumption intention, however, should show a negative relation (Xia, Monroe, and Cox 2004). As expected, our measurement correlates with all measures as hypothesized by literature (see table 3). Again, we correct correlations for the underlying measurement error.

5.1.6 Discriminant Validity

We assess whether our measurement is distinct from other linked constructs, namely price consciousness, price sentiment, price satisfaction, and demand for price regulations. We include multi-item measurements for the first three constructs and a single item measurement for the last construct. We refer to table 4 for the results of our discriminant validity analysis. We assess the average variances extracted of our multi-item measures by a confirmatory factor analysis for each construct and perceived inflation and determine the squared correlation between both (Fornell and Larcker 1981). Average variances extracted exceeds the squared correlations between the constructs in all cases. Moreover, squared correlation for our single-item measurement with perceived inflation is low.

Table 4: Discriminant Validity Measures

Measure	#Items	N	Coefficient Alpha	AVE	Squared Correlation with Perceived Inflation[1]
Perceived Inflation per Product Group	3	1,439	.949	.846	-
Price Sentiment	5	1,326	.823	.538	.319 ***
Price Consciousness	3	1,326	.737	.489	.003 *
Price Satisfaction	3	1,326	.846	.648	.286 ***
Demand for Price Regulation	1	1,326	-	-	.040 ***

1) correlations adjusted by reliability (see footnote 1); *$p < .1$, *$p < .05$, ***$p < .01$ (two-sided test).

5.2 Transforming Ratings into Percentages of Price Changes

5.2.1 Questionnaire

In the second step of our research approach, we identify a mapping function that translates scale values into percentage values. For this purpose, survey 2 consists of a quasi-experiment with 25 stimuli. Stimuli include a fictitious price change consisting of a current and a past price for a product in one of our 13 product groups. Appendix D provides details on our stimuli. Respondents rate each stimulus on our perceived inflation scale, which ensures consistency with survey 1. This approach enables us to match price changes with scale values, while respondents do not face any percentage values.

Overall, we include three survey versions. Two versions cover each four of our product groups and one version covers five groups. We vary stimuli for each version according to a reduced orthogonal design with factors product group, price of the product before price change (0 - 9.99 €, 10 € - 99.99 €, 100 - 999.99€), and size of the price change (-50% - -10%, -9.99% - 0%, 0.01% - 5%, 5.01% - 15%, 15.01% - 50%). We pretest survey 2 with a convenience sample of 206 respondents (86 female, 62 male, 58 respondents report no gender, 49 students).

5.2.2 Data

We collect data for survey 2 via the same way and time frame as survey 1. All respondents live in Germany, exceed the age of 16 years and have to pass a bogus item. Our sample includes 357 participants. 158 respondents state to be female (54.7%), 131 are male (45.3%). Mean age is 39.69 (SD 13.71) and mean net income is 2147.82 € (SD 1443.05). 118 respondents are university graduates (39.33%), 35 students (11.67%).

5.2.3 Model for our Mapping Function

Table E1 in appendix E summarizes the symbols and abbreviations of our variables and parameters. For our mapping function, we specify our model as:

$$PI_{s,k} = \alpha_k + \beta_k \cdot \pi_s + \varepsilon_{s,k} \tag{1}$$

with:

$$\alpha_k = \overline{\alpha} + \boldsymbol{\gamma}_\alpha^t \cdot \textbf{Sociodemographics}_k + \zeta_k \tag{2}$$

$$\beta_k = \overline{\beta} + \boldsymbol{\gamma}_\beta^t \cdot \textbf{Sociodemographics}_k + \upsilon_k \tag{3}$$

$$\varepsilon_{s,k}, \zeta_k, \upsilon_k \sim N(0; \Sigma) \,.$$

As, in our quasi-experiment, price changes are independent stimuli, we compute a percentage rate of inflation π_s for each stimulus s. Consumers rate these stimuli on our inflation scale. We, hence, regress average scale values $PI_{s,k}$ for respondents k and stimuli s on the respective percentage rate of inflation. α_k, β_k, γ_α^t, and γ_β^t are the parameters to be estimated and $\varepsilon_{s,k}$, ζ_k, and υ_k are error terms. In line with classical test theory, we assume equidistance for our scale and use a linear model.

We include interaction effects in the intercept and coefficient with sociodemographics. We identify gender, age, income, income change, individuals' involvement in daily shopping activities, household size, education, profession, and marital status as relevant sociodemographics from literature (Bates and Gabor 1986; Bruine de Bruin et al. 2011; Jonung 1981; Ranyard et al. 2008). These interactions allow us to construct an individual mapping function for each respondent. Apart from that, we include a random intercept and coefficient to control for additional heterogeneity not covered by the sociodemographics. For the transformation of rating values into percentage values, we use the inverse function of the model above.

5.2.4 Estimation and Findings

We estimate our model via a weighted maximum simulated likelihood approach with 1,000 random draws and freely correlated parameters. We weight observation according to respondent k's economic skills (ES) and mathematical skills (MS) by self-disclosure as below (for symbols and abbreviations see table E1 in appendix E):

$$w_k = \frac{ES_k + MS_k}{\sum_{k=1}^{K}(ES_k + MS_k)} \tag{4}$$

Table 5 shows the estimation results for our transformation function. They support our model choice. All sociodemographics reveal at least one significant interaction effect. However, interactions between intercept and marital status and relative price change and age, income, income change, and household size are insignificant at *p*>.1. In a holdout analysis, we show that our model outperforms other models. Appendix F provides details of this analysis.

Table 5: Estimation Results for the Mapping Function

Dependent variable: Average scale values for perceived inflation	Coefficients[1]		
Intercept	-.810	***	(.224)
SD	1.310	***	(.030)
x Sex (1=male)	.105	*	(.062)
x Age	.007	***	(.002)
x Income	$-.554\times10^{-4}$	**	$(.250\times10^{-4})$
x Income change	-.288	***	(.057)
x Individual's involvement into daily shopping activities	.101	***	(.024)
x Household size	.152	***	(.025)
x Education (1=graduate)	-.505	***	(.120)
x Profession (1=student)	-.364	***	(.063)
x Marital status (1=single)	.122		(.076)
Relative price change (stimulus)	9.119	***	(1.179)
SD	5.947	***	(.147)
x Sex (1=male)	-.733	**	(.329)
x Age	.010		(.013)
x Income	$.190\times10^{-3}$		$(.130\times10^{-3})$
x Income change	.080		(.302)
x Individual's involvement into daily shopping activities	-.403	***	(.130)
x Household size	-.055		(.134)
x Education (1=graduate)	1.503	**	(.596)
x Profession (1=student)	.740	**	(.333)
x Marital status (1=single)	1.420	***	(.409)
N	4962		
Pseudo-R^2	.632		
Log likelihood	-11934.117		

1) standard errors in parentheses; *p<.1, **p<.05, ***p<.01.

We use the inverse function of our model (equations (1)-(3)) and compute percentage values for our perceived inflation ratings. Table 6 shows official inflation rates, average perceived inflation rates and their difference per product group.

For 11 out of 13 product groups, perceived inflation exceeds official inflation. Overall, consumers tend to overestimate inflation. This result is in line with the majority of studies on inflation perceptions (e.g., Bates and Gabor 1986). But we see an adverse effect for the product groups alcoholic beverages and newspapers, books, and stationery. For three product groups, we even detect a perceived deflation in percentage figures. Concerning the absolute value of deviation of both figures, we see large variation across product groups. While alco-

holic beverages reveal only an absolute deviation of 1.37 percent points, perceived inflation exceeds official figures for transport services by 21.42 percent points.

Table 6: Descriptives for Official and Perceived Inflation

Product group	Official infla-tion rate [-∞,+∞][4]	Perceived inflation[1),3] [-5,+5][4]	Perceived infla-tion (trans-formed)[2),3] [-∞,+∞][4]	Difference per-ceived-official inflation [-∞,+∞][4]
Alcoholic beverages	1.32 %	-.061	-.05 %	-1.37 pp
Audio-visual, photographic and in-formation processing equipment	-2.77 %	-.585	-1.11 %	1.66 pp
Clothing	.28 %	.052	2.93 %	2.65 pp
Electricity, gas, and other fuels	-4.17 %	.864	13.24 %	17.41 pp
Household appliances	-1.68 %	-.253	.94 %	2.62 pp
Insurance services	.95 %	.501	8.48 %	7.53 pp
Medical products, appliances, and equipment	3.66 %	.702	9.15 %	5.49 pp
Newspapers, books, and stationery	1.63 %	-.411	-3.84 %	-5.47 pp
Operation of personal transport equipment	-3.43 %	.444	12.11 %	15.54 pp
Other recreational items and equip-ment, gardens and pet	.96 %	.689	13.70 %	12.74 pp
Personal care	1.36 %	-.004	3.95 %	2.59 pp
Recreational and cultural services	1.37 %	.534	8.36 %	6.99 pp
Transport services	3.57 %	1.616	24.99 %	21.42 pp

1) average scale value from our 3-item perceived inflation scale; 2) transformed by inverse function of equations (1)-(3); 3) weighted according to representativeness of observation for overall population in Germany; 4) values in parentheses represent the range that the specific measure is restricted to.

6 Driver Analysis

6.1 Data and Descriptive Results

For our driver analysis, we combine data from survey 1 and data on inflation by the German Federal Statistical Office (2016). Appendix C provides codings for data from survey 1.

6.1.1 Dependent Variables

We apply our transformation function to product class price changes. Based on these values, we construct absolute deviations of perceived and official price changes. In addition, we

determine weights in the perceived market basket via the following model (for symbols and abbreviations see table E1 in appendix E):

$$PI_{i,k} = \alpha_{i,k} + \sum_{j} \beta_{ij,k} \cdot PI_{ij,k} + \varepsilon_{i,k} \tag{5}$$

with:

$$\alpha_{i,k} = \overline{\alpha_i} + \zeta_{i,k} \tag{6}$$

$$\beta_{ij,k} = e^{\overline{\beta_{ij}} + \upsilon_{i,k}} \tag{7}$$

$$\varepsilon_{i,k}, \zeta_{i,k}, \nu_{ij,k} \sim N(0; \Sigma).$$

We regress consumer k's perceived inflation in a product group i ($P_{i,k}$) on her/his perceived inflation in the corresponding product classes ij ($P_{ij,k}$). $\alpha_{i,k}$ and $\beta_{ij,k}$ are the parameters that we estimate. $\varepsilon_{i,k}, \zeta_{i,k}$, and $\nu_{ij,k}$ are error terms. We use a random intercept and random parameter model. This model allows us to control for potential heterogeneity in market baskets across consumers. The random intercept controls for individual influences that arise from external factors not included in our analysis. Moreover, negative coefficients $\beta_{ij,k}$ would indicate that the perceived price development in a product class would negatively contribute to the price change perception for a corresponding product group. Therefore, we use a lognormal distribution for the parameters $\beta_{ij,k}$ that ensures that all $\beta_{ij,k}$ are positive. We use estimated $\beta_{ij,k}$ for the weights in a consumer's market basket.

We insert equations (6) and (7) into equation (5) and estimate a separate model per product group via a maximum simulated likelihood approach with 1,000 random draws. We allow parameters to be freely correlated. For reasons of representativeness, we weight observations according to their demographics. Table G1 in appendix G provides the detailed estimation results.

6.1.2 Drivers

We use data on official inflation to construct measures for the direction and size of price change, and the price change range. The direction of price change codes official price increases as 1, decreases as -1 and stability as 0 for each product class. The size of price change is the absolute value of the official inflation in a product class. For the price change range, we construct monthly inflation rates on official indexes between April 2014 and March 2015. Price change range is the difference between their maximum and minimum per product class.

Category clutter is the number of products in a product class according to the official market basket. Moreover, four raters individually code two dummy variables, which contain the value one when a product class contains mainly durables (dummy variable 1) or services (dummy variable 2) and zero otherwise. Initial interrater agreement by Krippendorff's alpha (Krippendorff 2013) based on 10,000 bootstraps is high with $\alpha_D = .911$ for durables and $\alpha_S = .974$ for services.

Table H1 in appendix H shows the descriptive values of our data for the driver analysis. As on the overall level, we see a general tendency to overestimate inflation ($M = 7.0$ percent points, SD = 32.8). However, we find high variation across individual values ranging from -79.5 percent points to 96.0 percent points. Large deviation also occurs for product classes. The highest mean deviation per product class results for passenger transport by railway with 37.7 percent points (SD = 25.5) and the lowest for games of chance with 18.9 percent points (SD = 21.0).

If the perceived inflation in a product group is fully explained by the perceived inflation in the corresponding product classes, all weights sum up to one. Hence, in groups with fewer classes the average weight per class is higher. We correct for this fact and multiply all weights with the number of product classes in a group. For full explanation, the average weight across groups would be one. Weights largely differ from official weights. The mean corrected

weight is .942 (SD = .723) and varies from .000 to 4.192 across individuals. Highest mean weight per product class evolves for electricity with 2.99 (SD = .436), lowest for beer with .078 (SD = .032).

Both, deviations of perceived and official inflation and perceived weights, show large variation. Similar variation exists across our drivers. For example, customer expertise is greatly higher for passenger transport by railway ($M = 4.550$, SD = 1.766) than games of chance ($M = 2.439$, SD = 1.699). For electricity, the monetary impact ($M = 187.604$, SD = 234.342) largely exceeds the one for beer ($M = 13.100$, SD = 10.536). These findings support our assumptions above.

6.2 Drivers of the Deviations of Perceived and Official Price Changes in Product Classes

6.2.1 Model of Drivers of Deviations of Perceived and Official Price Changes in Product Classes

We use a linear model as follows (for symbols and abbreviations see table E1 in appendix E):

$$\left|\pi_{ij} - \pi_{ij,k}^p\right| = \alpha_k + \beta_1 \cdot FRE_{ij,k} + \beta_2 \cdot REC_{ij,k} + \beta_3 \cdot DUR_{ij} + \beta_4 \cdot SER_{ij} + \beta_5 \cdot CC_{ij} +$$

$$\beta_6 \cdot PCR_{ij} + \beta_7 \cdot PPPF_{ij,k} + \beta_8 \cdot AOS_{ij,k} + \beta_9 \cdot PQR_{ij,k} + \beta_{10} \cdot PC_{ij,k} + \beta_{11} \cdot BL_{ij,k} + \beta_{12} \cdot$$

$$CE + \beta_{13} \cdot SA_{ij,k} + \beta_{14} \cdot NI_k + \beta_{15} \cdot ES_k + \beta_{16} \cdot SEX_k + \beta_{17} \cdot TSQ_k + \beta_{18} \cdot NPC_i + \varepsilon_{ij,k}$$

$$(8)$$

with

$$\alpha_k = \overline{\alpha} + \zeta_k \tag{9}$$

$$\varepsilon_{ij,k}, \zeta_k \sim N(0;\Sigma)$$

As we are interested in the deviation of official and perceived inflation rates, differences between these two rates per respondent k and product class j in product group i are our de-

pendent variables. Initially, both values are not comparable. Therefore, we construct percentage values for the perceived inflation in the product classes with our mapping function. We use absolute values for the differences because some drivers may lead to both, over- and underestimations. Absolute values ensure that these deviations do not cancel out. We regress them on the potential drivers. α_k and β's are the parameters to estimate. As every respondent rated the perceived inflation for three to six product classes, we control for this structure and include a respondent-specific random intercept.

6.2.2 Estimation Results

We insert equation (9) into equation (8) and estimate our model via a maximum simulated likelihood approach with 1,000 random draws. We allow free correlations between parameters. Table 7 shows our estimation results by ordinary least squares. Values are standardized coefficients from z-transformed independent variables. Our Pseudo-R^2 of .745 shows an excellent fit. Moreover, variance inflation factors are below the threshold of seven (Belsley 1991) and, thus, indicate no harmful multicollinearity.

6.2.3 Product Category Characteristics

Purchase frequency ($\beta_1 = .463$, $p = .019$) and recency ($\beta_2 = .657$, $p = .003$) display effects opposite to our previous expectations. Frequently and recently bought products often have low financial impact so that consumers pay low attention to prices and process them nonconsciously. Consumers may, hence, not store exact prices in memory (Monroe and Lee 1999). Moreover, consumers may displace older by more recent price information. Both issues may induce additional biases. The coefficients for nondurables ($\beta_3 = -594$, $p = .018$) and services ($\beta_4 = -1.149$, $p = .000$) support our assumptions. The effect of the category clutter ($\beta_5 = .318$, $p = .203$) shows the expected sign, but is insignificant. Price variability might be

high for an individual product, while on average different products in a product class show a similar price movement.

Table 7: Coefficients for the Driver Analysis

Dependent variable	Absolute deviation of official and perceived inflation in percent per product class		Weight per product class in the perceived market basket[2]	
	Standardized coefficients[1]			
Intercept	25.827***	(.200)	.967***	(.023)
SD	16.544***	(.199)	.000	(.018)
Product category characteristics				
Purchase frequency	.463**	(.197)	.049***	(.012)
Purchase recency	.657***	(.219)	.092***	(.015)
Product category (reference group: durables)				
Nondurables	-.594**	(.250)		
Services	-1.149***	(.266)		
Category clutter	.318	(.250)	.374***	(.022)
Monetary impact			.046***	(.015)
Price development				
Price change range	.254	(.192)	-.183***	(.025)
Perceived price promotion frequency	-.831***	(.216)	.079***	(.016)
Direction of price change			.056*	(.031)
Size of price change			.172***	(.035)
Official inflation rate			-.132***	(.032)
Consumer psychographics				
Affinity for online shopping	-.066	(.214)	-.001	(.015)
Price-quality relationship	-1.563***	(.198)		
Price consciousness	.494**	(.202)	-.006	(.016)
Brand loyalty	.348*	(.211)		
Customer expertise	1.937***	(.232)	-.018	(.015)
Economic environment				
Social Amplification	2.552***	(.204)	.023	(.016)
x Direction of price change			-.031	(.025)
Consumer demographics				
Income	-.934***	(.209)		
Economic skills	2.310***	(.212)		
Gender (reference group: female)	.012	(.205)		
Task characteristics				
Time spent on the questionnaire	1.211***	(.204)		
Number of product classes	1.428***	(.226)		
N / Pseudo-R²	3888 / .745		4996 / .234	

1) standardization results from a priori z-transformation of independent variables, standard errors in parentheses; 2) weights have been adjusted by multiplication by the number of product classes in the corresponding product group; *p<.1, **p<.05, ***p<.01.

48

6.2.4 Price Development

Price change range (β_6 = .254, p=.186) shows the expected positive sign, but is insignificant. Effects may be weaker on an aggregated level because a larger price change range may also help consumers to notice price changes. If consumers store past prices as a range, for small price changes, new prices may lie in this range so that consumers perceive no price change at all (Kalwani and Yim 1992; Ranyard et al. 2008). The coefficient for the perceived price promotion frequency (β_7 = -.832, p = .000) shows that promotions do not only support the development of a gut instinct on the accurate price, but the accurate price change, too.

6.2.5 Consumer Psychographics

The effect of affinity for online-shopping (β_8 = -.066, p = .760) does not reach significance. Lower search costs may face a higher information load online, which hampers consumers' price search. As expected, the involvement of the price-quality relationship (β_9 = -1.563, p = .000) stimulates price search and reduces the deviation of perceived and official inflation. The effects of price consciousness (β_{10} = .494, p = .015), brand loyalty (β_{11} = .348, p = .099), and customer expertise (β_{12}= 1.937, p = .000) support our assumptions and show a greater deviation of perceived and official inflation.

6.2.6 Economic Environment

Social amplification increases the deviation of perceived and official inflation (β_{13} = 2.552, p = .000). It reveals the largest coefficient of all effects. For our observation period, the official inflation rate was only .47%, while media reports are biased towards more extreme and rising prices. These facts may pronounce the influence of social amplification because media reports may reveal even higher biases than in times of higher inflation.

6.2.7 Controls

Gender is the only control variable that is insignificant (β_{16} = .012, p = .955). Traditional gender roles seem to be less relevant than in earlier studies (e.g., Jonung 1981).

6.3 Drivers of Weights in the Perceived Market Basket

6.3.1 Model of Drivers of Weights in the Perceived Market Basket

We use the following linear model (for symbols and abbreviations see table E1 in appendix A):

$$w_{ij,k} = \alpha_k + \beta_1 \cdot REC_{ij,k} + \beta_2 \cdot FRE_{ij,k} + \beta_3 \cdot CC_{ij} + \beta_4 \cdot MI_{ij,k} + \beta_5 \cdot PCR_{ij,k} + \beta_6 \cdot PPPF_{ij} + \beta_7 \cdot DPC_{ij} + \beta_8 \cdot SPC_{ij} + \beta_{7\times8} \cdot \pi_{ij} + \beta_9 \cdot PC_{ij,k} + \beta_{10} \cdot AOS_{ij,k} + \beta_{11} \cdot CE_{ij,k} + \beta_{12} \cdot SA_{ij,k} + \beta_{7\times12} \cdot DPC_{ij,k} \times SA_{ij,k} + \varepsilon_{ij,k}$$

$$(10)$$

with

$$\alpha_k = \overline{\alpha} + \zeta_k \tag{11}$$

$$\varepsilon_{ij,k}, \zeta_k \sim N(0; \Sigma)$$

Our dependent variable is a respondent k's weight of product class j in product group i in her/his perceived market basket. We correct weights for different numbers of product classes in a product group as explained above. We regress these weights on the potential drivers identified above and α_k and β's are the parameters to estimate. Again, we control for respondent-specific influences via a random intercept.

6.3.2 Estimation Results

We insert equation (10) into equation (11) and estimate our model via a weighted maximum simulated likelihood approach with 1,000 random draws. We allow free correlations between parameters. As our weights are results of a first-stage regression, we weight observa-

tions by their individual measurement error. You can find the estimation results in table 7. Values are standardized coefficients from z-transformed independent variables. Pseudo-R^2 of .234 indicates a good fit. Again, variance inflation factors indicate no multicollinearity.

6.3.3 Product Category Characteristics

As expected, purchase frequency (β_1 = .049, p = .000) and purchase recency (β_2 = .092, p = .000) increase the weight of a product class. The category clutter shows the largest effect (β_3 = .374, p = .000). In line with our assumptions, larger product diversity in a product class seems to enhance its salience and availability. Similarly, the monetary impact of a product class (β_4 = .046, p = .000) increases its weight.

6.3.4 Price Development

As expected, a larger price change range (β_5 = -.183, p = .000) decreases the comparability of price changes and, thus, the weight of that product class. Perceived price promotion frequency (β_6 =.079, p = .000) has a positive effect. Enhanced attention and more accurate processing of price information in frequently promoted categories actually seem to increase availability and salience. We find a positive influence of the direction of price change (β_7 = .056, p=.077) and the direction of price change (β_8 = .172, p = .000). Positive and larger absolute price changes seem, indeed, to increase salience and availability of a category. However, against our expectations, the coefficient for the official inflation rate ($\beta_{7\times8}$ = -.132, p = .000) is negative. The official inflation rate for April 2015 was .47%. Consumers might have perceived large and positive price changes as outliers which are less salient.

6.3.5 Consumer Psychographics

The effects of all consumer psychographics do not reach significance ($\beta_9 = -.001$, $p=.952$; $\beta_{10} = -.006$, $p = .721$; $\beta_{11} = -.018$, $p = .241$). It seems that they play a minor role for consumers' perceived market baskets.

6.3.6 Economic Environment

Contrary to our expectation, neither social amplification ($\beta_{10} = .023$, $p = .144$), nor its interaction with the direction of a price change ($\beta_{7 \times 12} = -.031$, $p = .206$) show a significant impact. Consumers seem to assume that reports from media and personal environment are biased. Hence, they do not attach additional importance to external press reports.

7 Discussion, Implications, Limitations, and Further Research

7.1 Discussion

Our study contributes to a meaningful measurement of perceived inflation. We develop and validate a tool that researchers, economists, and managers can use to track perceived inflation. Our scale reflects three mechanisms that companies use to change prices: changes in nominal prices, quantity, and quality of a product (Monroe 2003). We can, thus, show that consumers' perception of inflation reflects these mechanisms. Moreover, while previous studies almost exclusively find that consumers perceive higher inflation than official (e.g., Bates and Gabor 1986), we show that, depending on the product group, consumers may also underestimate inflation.

Our study gives insights on how perceived inflation develops at different aggregation levels. Deviations of higher-aggregated perceived inflation seem to arise in a two-step process. In a first step, consumers differ from official figures when they perceive individual price trends. In a second step, they further differ when they weight these for an aggregate value.

This result contributes to the research on bounded rationality because it shows that consumers' evaluation and aggregation of product price changes are subject to perceptional biases. However, as previous studies suggest, this process is far from being totally non-rational (e.g., Ranyard et al. 2008). In fact, we can show that consumers apply a process similar to the measurement of statistical offices.

While previous studies (e.g., Antonides 2008) have not distinguished these two steps, we analyze drivers for both steps in a separate analysis. We can show that it is important to distinguish here. Against existing concepts (Ranyard et al. 2008), social amplification influences deviations of price changes, but not the individual aggregation process. Moreover, effects evolve more complex when researchers differentiate between both steps. For example, perceived price promotion frequency reduces the deviation of perceived and official inflation, but increases the weight of a product class. Contrary, a higher purchase recency and frequency lead to a higher deviation of perceived and official inflation in a product class and a higher weight.

Our results contribute to research on price recall accuracy (e.g., Vanhuele and Drèze 2002). We find evidence that a better knowledge of single prices may not lead to a better perception of price changes per se. If recall accuracy of single prices, like for price conscious consumers, leads to lower assessment of all other prices in a product class, it causes larger deviation of perceived and official inflation in the overall product class. In this case, correct knowledge of a product price may even hamper the accurate perception of higher-aggregated inflation. Hence, higher price recall accuracy may not necessarily induce higher accuracy for the perception of price trends.

In addition, we find that customer expertise relates negatively to the accuracy of perceived inflation for product classes, while using the price as a quality signal leads to the oppo-

site effect. We, hence, contribute to the research on subjective knowledge (e.g., Alba and Hutchinson 2000) which shows that overconfidence may lead to inferior evaluations.

While previous studies have predominantly related large and positive price changes to higher-aggregated perceived inflation (e.g., Bruine de Bruin, van der Klaauw, and Topa 2011), we show that other factors come into play, when consumers weight price trends. In line with Lourenço, Gijsbrechts, and Paap (2015), availability, salience, and comparability are those factors that determine these weights. Hence, although positive and larger price changes each have a higher influence on the weight in the perceived market basket, this effect is reduced for large and positive price changes. We lead this effect back to the overall low inflation rate and, thus, reduced salience and comparability of high price changes at the time of our study. However, drivers also differ from the perception of aggregated price images. While consumers are, here, relatively insensitive concerning price promotions (Lourenço, Gijsbrechts, and Paap 2015), perceived price promotion frequency increases the weight in the perceived market basket for a product class.

7.2 Implications

Retail managers may use our results to restructure and plan their portfolios. As high perceived inflation leads to lower consumption, they may offer less or cheaper products in those categories with high-perceived inflation. Moreover, as media reports are a main driver of the deviation of perceived and official inflation on a product level, they may use media reports as an indicator of future consumption shifts.

Likewise, brand managers may use perceived inflation to plan their pricing and promotion strategies. In particular, categories with high relative importance need a detailed planning. These categories might be categories with a high degree of diversification or a more recent and frequent purchase pattern. One way to increase prices might here be to combine

increases with large, but rare price promotions. When promotion frequency is low, consumers may not precisely track price trends and even attach a lower weight to a category.

The results of our study may help economists to reach a desired consumption or saving rate. They may apply (de)regulations to those categories that have a higher importance for consumers' perceived inflation. These categories are more diversified, have been frequently and recently bought and come with a larger amount of money per purchase.

7.3 Limitations

We only measure perceived inflation for Germany. However, perceived inflation and its deviation from official values may vary across countries (Bates and Gabor 1986; Jonung 1981). It is up to future research to analyze how cultural factors influence the formation of perceived inflation. Finally, we use a selection of product groups and apply their composition as given by the German Federal Statistical Office (2016). Other product groups might reveal different perception processes and consumers might perceive other compositions.

REFERENCES PROJECT 1

Alba, Joseph W., and J. Wesley Hutchinson (2000). "Knowledge Calibration: What Consumers Know and What They Think They Know," *Journal of Consumer Research*, 27 (2), 123-156.

Antonides, Gerrit (2008). "How Is Inflation Related to Actual Price Changes in the European Union?," *Journal of Economic Psychology*, 29 (4), 417-432.

Atkinson, Richard C. and Richard M. Shiffrin (1968). "Human Memory: A Proposed System and Its Control Processes," in K. W. Spence, J. T. Spence (Eds.), *The Psychology of Learning and Motivation*, New York: Academic Press, 89-195.

Bates, John M. and André Gabor (1986). "Price Perception in Creeping Inflation: Report on an Enquiry," *Journal* of Economic Psychology, 7 (3), 291-314.

Belsley, David A. (1991). *Conditioning Diagnostics: Collinearity and Weak Data Regression*, New York: Wiley.

Bijmolt, Tammo H. A., Harald J. van Heerde, and Rik G. M. Pieters (2005). "New Empirical Generalizations on the Determinants of Price Elasticity," *Journal of Marketing Research*, 42(2), 141-156.

Bolton, Lisa E., Luk Warlop, and Joseph W. Alba (2003). "Consumer Perceptions of Price (Un)Fairness," *Journal of Consumer Research*, 29 (4), 474-491.

Brachinger, Wolfgang (2008). "A New Index of Perceived Inflation: Assumptions, Method and Application to Germany," *Journal of Economic Psychology*, 29 (4), 433-457.

Bruine de Bruin, Wändi, Charles F. Manski, Giorgio Topa and Wilbert van der Klaauw (2011). "Measuring Consumer Uncertainty about Future Inflation," *Journal of Applied Econometrics*, 26 (3), 454-478.

Bruine de Bruin, Wändi, Wilbert van der Klaauw, and Giorgio Topa (2011). "Expectations of Inflation: The Biasing Effect of Thoughts about Specific Prices," *Journal of Economic Psychology*, 32 (5), 834-45.

Bureau of Labor Statistics (2016). *Consumer Price Index* [available at http://www.bls.gov/cpi/cpiovrvw.htm#item2; accessed December 13, 2016].

Chen, Haipeng (Allan) and Akshay R. Rao (2007). "When Two Plus Two Is Not Equal to Four: Errors in Processing Multiple Percentage Changes," *Journal of Consumer Research*, 34 (3), 327-340.

Desai, Kalpesh K. and Debabrata Talukdar (2003). "Relationship between Product Group's Price Perceptions, Shopper's Basket Size, and Grocery Store's Overall Store Price Image," *Psychology and Marketing*, 20 (10), 903-933.

DeVellis, Robert (2012). *Scale Development*, Los Angeles et al.: Sage.

Dräger, Lena, Jan-Oliver Menz, and Ulrich Fritsche (2014). "Perceived Inflation under Loss Aversion," *Applied Economics*, 46 (3), 282-293.

Estelami, Hooman and Donald R. Lehmann (2001). "The Impact of Research Design on Consumer Price Recall: An Integrative View," *Journal of the Academy of Marketing Science*, 29 (1), 36-49.

Estelami, Hooman and Peter de Maeyer (2004). "Product Category Determinants of Price Knowledge for Durable Goods," *Journal of Retailing*, 80 (2), 129-137.

EU Consumer Survey (2016). *Questionnaires in National Languages* [available at http://ec.europa.eu/economy_finance/db_indicators/surveys/questionnaires/index_en.htm; accessed December 13, 2015].

Fisher, Irving (1922). *The Making of Index Numbers: A Study of Their Varieties, Tests, and Reliability*, Boston et al.: Houghton Mifflin.

Fornell, Claes and David F. Larcker (1981). "Evaluating Structural Equation Models with Unobservable Variables and Measurement Error," *Journal of Marketing Research*, 18 (1), 39-50.

German Federal Statistical Office (2016). *Consumer Price Indices* [available at http://www.destatis.de; accessed December 13, 2016].

Jonung, Lars (1981). "Perceived and Expected Rates of Inflation in Sweden," *American Economic Review*, 71 (5), 960-968.

Jonung, Lars (1986). "Uncertainty about Inflationary Perceptions and Expectations," *Journal of Economic Psychology*, 7 (3), 315-325.

Jonung, Lars and David Laidler (1988). "Are Perceptions of Inflation Rational? Some Evidence for Sweden," *American Economic Review*, 78 (5), 1080-1087.

Kahneman, Daniel and Amos Tversky (1972). "Subjective Probability: A Judgment of Representativeness," *Cognitive Psychology*, 3 (3), 430-454.

Kalwani, Manohar U. and Chi Kin Yim (1992). "Consumer Price and Promotion Expectations: An Experimental Study," *Journal of Marketing Research*, 29 (1), 90-100.

Katona, George (1974). "Psychology and Consumer Economics," *Journal of Consumer Research*, 1 (1), 1-8.

Krippendorff, Klaus (2013). *Content Analysis*, Los Angeles et al.: Sage.

Lichtenstein, Donald R., Nancy M. Ridgway, and Richard G. Netemeyer (1993). "Price Perceptions and Consumer Shopping Behavior: A Field Study," *Journal of Marketing Research*, 30 (2), 234-245.

Lourenço, Carlos J. S., Els Gijsbrechts, and Richard Paap (2015). "The Impact of Category Prices on Store Price Image Formation: An Empirical Analysis," *Journal of Marketing Research*, 52 (2), 200-216.

Lynch, John G. and Dan Ariely (2000). "Wine Online: Search Cost Affect Competition on Price Quality and Distribution," *Marketing Science*, 19 (1), 83-103.

Monroe, Kent B. (2003). *Pricing: Making Profitable Decisions*, Burr Ridge: McGraw-Hill.

Monroe, Kent B. and Angela Y. Lee (1999). "Remembering versus Knowing: Issues in Buyers' Processing of Price Information," *Journal of the Academy of Marketing Science*, 27 (2), 207-225.

Ranyard, Rob, Fabio del Missier, Nicolao Bonini, Darren Duxbury and Barbara Summers (2008). "Perceptions and Expectations of Price Changes and Inflation: A Review and Conceptual Framework," *Journal of Economic Psychology*, 29 (4), 378-400.

Soroka, Stuart N. (2006). "Good News and Bad News: Asymmetric Responses to Economic Information," *Journal of Politics*, 68 (2), 372-385.

Tversky, Amos, and Daniel Kahneman (1973). "Availability: A Heuristic for Judging Frequency and Probability," *Cognitive Psychology*, 5 (2), 207-232.

United Nations Statistics Division (2016). *Classification of Individual Consumption According to Purpose* [available at http://unstats.un.org/unsd/cr/registry/regcst.asp?Cl=5; accessed December 13, 2016].

Vanhuele, Marc and Xavier Drèze (2002). "Measuring the Price Knowledge Shoppers Bring to the Store," *Journal of Marketing*, 66 (4), 72-85.

Xia, Lan, Kent B. Monroe and Jennifer L. Cox (2004). "This Price is Unfair! A Conceptual Framework of Price Fairness Perceptions," *Journal of Marketing*, 68 (4), 1-15.

Appendix A: Product Categories and Groups in our Sample

We measure perceived inflation across 13 product groups and corresponding product classes. We take these groups and classes from the Classification of Individuals Consumption According to Purpose (COICOP, for details see United Nations Statistics Division (2016)). Thereby, groups consist of three to six corresponding product classes, which sum up totally to 53 product classes. Table W1 below provides the detailed structure of the product groups and classes that we use for our analysis. When we refer to the numeration of product groups and classes in the subsequent materials, we use numerations as given in table A1.

Table A1: Overview on Product Groups and Classes

Product Group		Product Class	
No.	*Name*	*No.*	*Name*
1	Alcoholic beverages	1	Spirits
		2	Wine
		3	Beer
		4	Mixed drinks (<6% alcoholic strength)
2	Audio-visual, photographic and information processing equipment	1	Equipment for the reception, recording and reproduction of sound and pictures
		2	Photographic and cinematographic equipment and optical instruments
		3	Information processing equipment
		4	Recording media
		5	Repair of audio-visual, photographic and information processing equipment
3	Clothing	1	Clothing materials
		2	Garments
		3	Other articles for clothing and clothing accessories
		4	Cleaning, repair and hire of clothing
4	Electricity, gas, and other fuels	1	Electricity
		2	Gas
		3	Liquid fuels
		4	Solid fuels
		5	Heat energy
5	Household appliances	1	Major household appliances whether electric or not
		2	Small electric household appliances
		3	Repair services for household appliances

Table A1: Overview on Product Groups and Classes (Cont'd)

Product Group		Product Class	
No.	*Name*	*No.*	*Name*
6	Insurance services	1	Insurance connected with the dwelling
		2	Insurances connected with health
		3	Insurance connected with transport
		4	Other insurance
7	Medical products, appliances, and equipment	1	Pharmaceutical products
		2	Other medical products
		3	Therapeutically appliances and equipment
8	Newspapers, books, and stationery	1	Books
		2	Newspapers and periodicals
		3	Miscellaneous printed matter
		4	Stationery and drawing materials
9	Operation of personal transport equipment	1	Spare Parts and accessories for personal transport equipment
		2	Fuels and lubricants for personal transport equipment
		3	Maintenance and repair of personal transport equipment
		4	Other services of personal transport equipment
10	Other recreational items and equipment, gardens and pets	1	Games, toys, and hobbies
		2	Equipment for sport, camping and open-air recreation
		3	Gardens, plants and flowers
		4	Pets and related products
		5	Veterinary and other services for pets
11	Personal care	1	Hairdressing salons and personal grooming establishments
		2	Electric appliances for personal care
		3	Other appliances, articles and products for personal care
12	Recreational and cultural services	1	Recreational and sporting services
		2	Cultural services
		3	Games of chance
13	Transport services	1	Passenger transport by railway
		2	Passenger transport by road
		3	Passenger transport by air
		4	Passenger transport by sea and inland waterway
		5	Combined passenger transport
		6	Other purchased transport services

Appendix B: Details on Item Generation and Scale Purification

For the scale development, we follow a four-step approach. First, we generate items. Our aim is to cover all potential aspects of price changes. Hence, we combine several sources. We construct items based on the result of a pre-study, other measures of perceived inflation, and related literature from various areas, such as price management (e.g., Monroe 2003). Overall, we create an item battery of 158 items.

Second, we focus on the content and face validity of our items. Therefore, we discuss and evaluate items with the aid of potential respondents and experts in marketing. They rate our items concerning clarity, relevance, specificity, and representativeness. We take these ratings and discuss, adapt, and exclude items. 37 items remain.

Third, we apply these items together with several validity scales in a first pretest for three product groups. We use a convenience sample with 118 respondents. Among this sample, 63 respondents state to be female, 52 to be male (2 respondents do not provide their gender), and 90 respondents are students. We reduce items considering reliability measures and correlations with validity measures. We exclude items so that coefficient alpha still exceeds .85 and average scale values show meaningful correlations to validity measures. Six items result.

Based on these results, we conduct another pretest including the overall questionnaire of survey 1 for four product groups. Our sample is a convenience sample with 224 respondents. From these respondents, 78 state to be female, 64 state to be male (82 respondents do not provide their gender), and 100 are students. We reduce the scale to our final three items by using reliability measures and correlations with validity measures as above.

Appendix C: Survey Measures, Translations, and Origin

We conducted surveys in German.) below provides translations of all scales that we use including their origins. In addition, table C2 provides the original German measures.

Table C1: Translations of Measures within the Questionnaire[1]

Perceived inflation[3]
  Prices have risen strongly within the last 12 months.
  The same consumption of products and services costs significantly more today than 12 months ago.
  Today, I pay substantially more for the same quality than 12 months ago.

Convergent validity measures
  Price tendency (EU Consumer Survey 2016)
      How do you think that consumer prices in the category [category] have developed over the last 12 months? They have… (-1 = fallen; 0 = stayed about the same, 1 = risen slightly/moderately/a lot)
  Percentage statement (EU Consumer Survey 2016)
      By how many per cent do you think prices in the category [category] have gone up/down over the last 12 months?

Nomological validity measures
  Expected inflation
      Prices will rise strongly within the next 12 months.
      The same consumption of products and services will cost significantly more in 12 months compared to today.
      In 12 months, I will pay substantially more for the same quality than today.
  Consumption intention[4] (Kozup, Creyer, and Burton 2003; Burton, Garretson, and Veliquette 2012)
      Given the price development over the last 12 months, would you buy more products and services in [category]?[2] (1 = more, …, 7 = less)
      According to the price development over the last 12 months, how likely is it that you consider buying more products and services in [category]?[2] (1 = very likely, …, 7 = unlikely)
      How realistic is it that you buy more products and services in [category], given the price development over the last 12 months?[2] (1 = very realistic, …, 7 = not realistic)
  Saving intention (Yamamuchi and Templer 1982; Strutton and Lewin 2012)
      One should currently try to save money otherwise spent for products and services in [category].
      One should currently try to spend less on products and services in [category].
      One should currently try to avoid excessive debts for products and services in [category].
      One should currently spend more money on products and services in other categories than spending money on products and services in [category].
  Actual postponement of purchases[4] (Strutton and Lewin 2012)
      Over the last 12 months, I have postponed purchasing products and services in [category].
  Intended postponement of purchases[4]
      Currently, one should wait a little to purchase products and services in [category].

Table C1: Translations of Measures within the Questionnaire[1] (Cont'd)

Price fairness[3] (Drake and Dahl 2003)

   The prices over the last 12 months have been fair.
   The prices over the last 12 months have been justified.
   The prices over the last 12 months have been honest.
   The prices over the last 12 months have been unfair. [2]
   The prices over the last 12 months have been questionable. [2]
   I have perceived prices over the last 12 months as a "rip-off".[2]

Discriminant validity measures

Price sentiment[4] (Gaski and Etzel 1986)

   Most products I have bought over the last 12 months have been overpriced.
   Businesses could have charged lower prices over the last 12 months and still been profitable.
   Most prices over the last 12 months have been reasonable.[2]
   Most prices over the last 12 months have been fair.
   In general, I am satisfied with the prices I have paid over the last 12 months.

Price consciousness (multi-item)[3] (Darden and Perreault 1976)

   I compare prices for at least a few providers before I choose one.
   I find myself checking the prices even for small things.
   It is important to me to get the best prices for the products and services I buy.

Price satisfaction[3] (Homburg, Koschate-Fischer, and Wiegner 2012)

   The prices over the last 12 months have totally met my expectations.
   The prices over the last 12 months have been almost ideal.
   All in all, I have been very satisfied with the prices over the last 12 months.

Demand for price regulation[3] (Lundstrom and Lamont 1976)

   Permanent price controls are the only way to stop price increases.

Weighting factors

Economic Skills

   How would you evaluate your economic skills? (1: very bad, …, 7: very good)

Mathematical Skills

   How would you evaluate your mathematical skills? (1: very bad, …, 7: very good)

Substantive drivers

Purchase frequency

   How often did you buy [category] on average? (3 = more than once a week, 1 = once a week, .5 = once every 14 days, .25 = once a months, .125 = once to twice every three months, .02 = less frequent)

Purchase recency

   When have you bought [category] the last time? (-1.5 = within the last three days, -6.5 = within the last 14 days, -39 = within the last two months, -212.5 = within the last 12 months, -730= more than 12 months ago)

Perceived price promotion frequency[4]

   In [category], there are often special offers.

Price consciousness (Single-Item)[4] (Darden and Perreault 1976)

   I find myself checking prices even for small things in [category].

Affinity for online shopping[3]

   I predominantly buy [category] online.

Table C1: Translations of Measures within the Questionnaires[1] (Cont'd)

Brand loyalty (Ailawadi, Neslin, and Gedenk 2001)

    For most of the products and services from the category [category] I buy, I have a provider which I prefer.

Customer expertise[4] (Chang 2004)

    I know a lot about [category].

Price-quality relationship[4] (Estelami and de Maeyer 2004)

    The higher the price for [category], the higher the quality.

Social amplification

    How frequently have you noticed reports about price changes in [category] in media within the last 12 months? (1=hardly, …, 7=frequently)

    How frequently have you noticed reports about price changes in [category] from your personal environment within the last 12 months? (1=hardly, …, 7=frequently)

Monetary impact

    When you buy [category], how much do you spend on average? (5 = 0.00 - 9.99€, 20 = 10.00 - 29.99€, 65 = 30.00 – 99.99 €, 175 = 100 – 249.99 €, 375 = 250.00 – 500.00€, 750 = more than 500.00 €)

---

1) names in parentheses behind scale names represent origin of scale items that we have adapted for our survey, coding is given in parentheses behind items, for Likert scales, scale values represent codings, original scale items were German; 2) reversed item; 3) 11-point Likert scale (-5: I do not agree at all, …, 5: I totally agree); 4) 7-point Likert scale (1: I do not agree at all, …, 5: I totally agree).

Table C2: Original Measures of the Questionnaires in German[1]

Perceived inflation[3]

    Die Preise sind in den vergangenen 12 Monaten stark gestiegen.

    Der gleiche Verbrauch an Produkten und Leistungen kostet mich heute wesentlich mehr als vor 12 Monaten.

    Ich zahle heute wesentlich mehr für die gleiche Qualität als vor 12 Monaten.

Convergent validity measures

Price tendency

    Wie haben sich Ihrer Ansicht nach die Verbraucherpreise in der Kategorie [category] in den letzten 12 Monaten entwickelt?

    Sie sind… (-1 = gesunken; 0 = in etwa gleich geblieben, 1 = leicht/in Maßen, stark gestiegen)

Percentage statement

    Um wie viel Prozent sind Ihrer Ansicht nach die Verbraucherpreise in der Kategorie [category] in den letzten 12 Monaten gesunken/gestiegen?

Nomological validity measures

Expected inflation

    Die Preise werden in den nächsten 12 Monaten stark steigen.

    Der gleiche Verbrauch an Produkten und Leistungen wird mich in 12 Monaten wesentlich mehr kosten als heute.

    Ich werde in 12 Monaten wesentlich mehr für die gleiche Qualität zahlen müssen als heute.

Table C2: Original Measures of the Questionnaires in German [1] (Cont'd)

Consumption intention[4]

Würden Sie, entsprechend der Preisentwicklung in den letzten 12 Monaten, jetzt mehr Produkte und Leistungen im Bereich [category] kaufen?[2] (1 = mehr, …, 7 = weniger)

Entsprechend der Preisentwicklung in den letzten 12 Monaten, wie wahrscheinlich ist es, dass Sie es jetzt in Erwägung ziehen, mehr Produkte und Leistungen im Bereich [category] zu kaufen??[2] (1 = sehr wahrscheinlich, …, 7 = nicht wahrscheinlich)

Wie realistisch ist es, dass Sie jetzt mehr Produkte und Leistungen im Bereich [category] kaufen, entsprechend der Preisentwicklung in den letzten 12 Monaten?[2] (1 = sehr realistisch, …, 7 = nicht realistisch)

Saving intention

Man sollte aktuell versuchen bei Produkten und Leistungen im Bereich [category] mehr zu sparen.

Man sollte aktuell versuchen bei Produkten und Leistungen im Bereich [category] weniger auszugeben.

Man sollte aktuell versuchen Kredite für Produkte und Leistungen im Bereich [category] zu vermeiden.

Man sollte aktuell eher Geld für Produkte und Leistungen in anderen Kategorien ausgeben, anstatt Geld für Produkte und Leistungen im Bereich [category] auszugeben.

Actual postponement of purchases[4]

In den letzten 12 Monaten habe ich den Erwerb von Produkten und Leistungen im Bereich [category] auf später verschoben.

Intended postponement of purchases[4]

Man sollte aktuell mit dem Erwerb von Produkten und Leistungen im Bereich [category] noch etwas warten.

Price fairness[3]

Die Preise in den letzten 12 Monaten waren fair.

Die Preise in den letzten 12 Montan waren gerechtfertigt.

Die Preise in den letzten 12 Montan waren ehrlich.

Die Preise in den letzten 12 Montan waren unfair.[2]

Die Preise in den letzten 12 Montan waren fragwürdig.[2]

Ich habe die Preise in den letzten 12 Monaten als "Abzocke" empfunden.[2]

Discriminant validity measures

Price sentiment[4]

Die meisten Produkte, die ich in den letzten 12 Monaten gekauft habe, waren überteuert.

Unternehmen hätten in den letzten 12 Monaten günstigere Preise verlangen können und hätten immer noch Gewinne gemacht.

Die meisten Preise in den letzten 12 Monaten waren angemessen.[2]

Die meisten Preise in den letzten 12 Monaten waren fair.

Generell bin ich zufrieden mit den Preisen, die ich in den letzten 12 Monaten gezahlt habe.

Price consciousness (multi-item)[3]

Ich vergleiche zumindest die Preise einiger Anbieter, bevor ich einen Anbieter wähle.

Ich vergleiche Preise selbst für kleine Sachen.

Es ist für mich wichtig, die besten Preise für Produkte und Leistungen zu erhalten, die ich erwerbe.

Price satisfaction[3]

Die Preise trafen in den letzten 12 Monaten genau meine Erwartungen.

Die Preise waren nahezu ideal in den letzten 12 Monaten.

Alles in allem, war ich in den letzten 12 Monaten mit den Preisen sehr zufrieden.

Demand for price regulation[3]

Permanente Preiskontrollen sind der einzige Weg, um Preissteigerungen zu stoppen.

Table C2: Original Measures of the Questionnaires in German [1] (Cont'd)

Weighting factors
Economic Skills
  Wie beurteilen Sie Ihre wirtschaftlichen Kenntnisse? (1: sehr schlecht, …, 7: sehr gut)
Mathematical Skills
  Wie beurteilen Sie Ihre mathematischen Kenntnisse? (1: sehr schlecht, …, 7: sehr gut)

Substantive drivers
Purchase frequency
  Wann oft kaufen Sie [category] im Durchschnitt? (3 = täglich, 1 = einmal pro Woche, .5 = einmal alle 14 Tage, .25 = einmal im Monat, .125 = einmal bis zweimal alle drei Monate, .02 = seltener)
Purchase recency
  Wann haben Sie [category] zum letzten Mal gekauft? (-1.5 = innerhalb der letzten drei Tage, -6.5 = innerhalb der letzten 14 Tage, -39 = innerhalb der letzten zwei Monate, -212.5 = innerhalb der letzten 12 Monate, -730= vor mehr als 12 Monaten)
Perceived price promotion frequency[4]
   Bei [category] gibt es häufig Sonderangebote.
Price consciousness (Single-Item)[4]
  Ich vergleiche die Preise selbst für kleine Dinge im Bereich [category].
Affinity for online shopping[3]
  Ich kaufe [category] vorwiegend online.
Brand loyalty
  Für die meisten Produkte und Leistungen aus dem Bereich [category], die ich kaufe, habe ich einen Anbieter, den ich präferiere.
Customer expertise[4]
  Ich kenne mich gut mit [category] aus.
Price-quality relationship[4]
  Je höher der Preis für [category], desto höher die Qualität.
Social amplification
  Wie stark haben Sie in den letzten 12 Monaten Berichte über Preisänderungen im Bereich [category] in den Medien wahrgenommen? (1=kaum, …, 7=häufig)
  Wie stark haben Sie in den letzten 12 Monaten Berichte über Preisänderungen im Bereich [category] aus Ihrem Umfeld vernommen? (1=kaum, …, 7=häufig)
Monetary impact
  Wenn Sie [category] kaufen, wie viel geben Sie bei einem durchschnittlichen Kauf hierfür aus? (5 = 0.00 - 9.99€, 20 = 10.00 - 29.99€, 65 = 30.00 – 99.99 €, 175 = 100 – 249.99 €, 375 = 250.00 – 500.00€, 750 = mehr als 500.00 €)

1) for Likert scales, scale values represent codings, depending on the category, item wording may vary slightly; 2) reversed item; 3) 11-point Likert scale (-5: Stimme ganz und gar nicht zu, …, 5: Stimme voll und ganz zu); 4) 7-point Likert scale (1: Stimme ganz und gar nicht zu, …, 5: Stimme voll und ganz zu).

Appendix D: Stimuli of the Quasi-Experiment

For survey 3, we use 3 different versions. Each version compromises of 25 stimuli. Within the versions, we randomize the order of the stimuli across respondents. We build all stimuli according to the same pattern. The basic structure of the stimuli appears as follows:

[product] was [former price] € 12 months ago. The average price has [direction of price change] by [price change] € to [new price] €. Prices for other goods and services from the category [product group] have developed correspondingly. If you consider this price change, how would you evaluate the following statements concerning the category [product group]?

We substituted parts in brackets according to the following table D1:

Table D1: Stimuli in Survey 2

Version	Stimulus	Product group	Product	Former Price	Direction of price change	Price Change	New Price
1	1	audio-visual, photographic and information processing equipment	an unrecorded DVD	1.13	fallen	0.15	0.98
	2				risen	0.01	1.14
	3		a pair of headphones	99.00	fallen	6.87	92.13
	4				risen	13.63	112.63
	5		television set	899.00	risen	222.59	1121.59
	6	insurance services	daily fee for a liability insurance	0.53	fallen	0.01	0.52
	7					0.22	0.31
	8				risen	0.06	0.59
	9					0.22	0.75
	10		monthly fee for a supplementary health insurance	37.50	fallen	1.88	35.62
	11					17.06	20.44
	12				risen	1.55	39.05
	13					18.00	55.50
	14		yearly fee for an accident insurance	385.25	risen	12.54	397.79
	15				risen	56.78	442.03
	16	newspapers, books, and stationery	paperback	8.21	risen	0.32	8.53
	17					3.84	12.05
	18		stylograph	26.50	fallen	3.70	22.80
	19				risen	3.95	30.45
	20		subscription fee for a daily newspaper per year	358.80	fallen	7.59	351.21

Table D1: Stimuli in Survey 2 (Cont'd)

1	21	recreational and cultural service	visit to a swimming pool	4.88	fallen	0.46	4.42
	22				risen	0.51	5.39
	23		visit to a theme park	44.66	risen	0.92	45.58
	24					13.89	58.55
	25		annual season ticket of a soccer club	638.26	fallen	250.25	388.01
2	26	clothing	scarf	6.48	risen	0.39	6.87
	27					2.98	9.46
	28		shirt	49.21	fallen	0.14	49.07
	29				risen	1.66	50.87
	30		leather jacket	389.78	fallen	104.48	285.30
	31	electricity, gas, and other fuels	1 kWh electricity	0.25	fallen	0.01	0.24
	32				changed	0.00	0.25
	33		bottle of propane gas	33.85	fallen	3.71	30.14
	34				risen	11.12	44.97
	35		tankful (1,000 l) of domestic fuel oil	777.52	risen	57.03	834.55
	36	household appliances	egg boiler	7.22	fallen	0.96	6.26
	37				risen	0.32	7.54
	38					0.44	7.66
	39					1.10	8.32
	40		water boiler	26.38	fallen	13.10	13.28
	41					1.95	24.43
	42				risen	0.55	26.93
	43					1.46	27.84
	44		refrigerator	459.63	fallen	23.71	435.92
	45				risen	99.55	559.18
	46	operation of personal transport equipment	1 l gas	1.55	fallen	0.03	1.52
	47					0.64	0.91
	48		bicycle tire	82.27	risen	6.39	88.66
	49					26.93	109.20
	50		car repair	850.00	risen	14.99	864.99
3	51	alcoholic beverages	a bottle of beer (0.33 l)	1.03	fallen	0.16	0.87
	52				risen	0.09	1.12
	53		a bottle of whisky (0.75 l)	51.21	fallen	4.06	47.15
	54					9.33	60.54
	55		a bottle of champagne (0.75 l)	143.98	risen	5.55	149.53
	56	medical products, appliances, and equipment	a package of aspirin (40 tablets)	8.52	fallen	1.09	7.43
	57				risen	0.27	8.79
	58		medical thermometer	23.18	fallen	0.75	22.43
	59				risen	2.99	26.17
	60		progressive power lens	324.60	risen	75.68	400.28
	61	Other recreational items and equipment, gardens and pets	a bag of potting soil (5 l)	5.51	fallen	0.01	5.50
	62				risen	0.57	6.08
	63		a puppet	67.47	risen	0.38	67.58
	64					14.15	81.62
	65		a pair of skis	558.08	fallen	118.69	439.39

Table D1: Stimuli in Survey 2 (Cont'd)

3	66	personal care	a pack of toilet	3.74	risen	0.09	3.83
	67		paper (8 rolls)			0.77	4.51
	68		a woman's haircut	59.22	fallen	16.89	42.33
	69				risen	3.69	62.91
	70		a bottle of perfume (200 ml)	182.31	fallen	16.75	165.56
	71	transportation	a single ride	2.31	fallen	0.04	2.27
	72	services	(local traffic)		risen	0.35	2.66
	73		a weekly ticket	28.84	fallen	10.17	18.67
	74		(local travel)		risen	0.84	29.68
	75		a flight ticket (long-distance)	983.36	risen	100.01	1083.37

For example, if we choose stimulus 31, the stimulus in the survey appeared as follows (substituted parts are italicized):

1 kWh electricity was *0.25 €* 12 months ago. The average price has *fallen* by *0.01 €* to *0.24 €*. Prices for other goods and services from the category *electricity, gas, and other fuels* have developed similarly. If you consider this price change, how would you evaluate the following statements concerning the category *electricity, gas, and other fuels*?

The structure of the stimuli results from an orthogonal design for each version across three factors. Underlying factors are the product group, the level of the original price, and the size of the price change.

For the level of the original price we chose 0 - 9.99 €, 10 € - 99.99 €, and 100 - 999.99€. We choose these stages to cover a broad range of price change scenarios. For each product group, we choose one product that fits into the given price range for the German market. We work with products instead of product groups or classes. They are easier to comprehend and represent a more realistic scenario than abstract price changes for higher aggregation levels.

For the price changes we choose the following ranges -50% - -10%, -9.99% - 0%, 0.01% - 5%, 5.01% - 15%, and 15.01% - 50%. Within the ranges we draw a random price change for

each stimulus. These ranges match the distribution of real inflation rates in recent years for Germany as given in figure D1. Figure D1 shows annual inflation rates between March 2010 and March 2015 across all product classes that the German statistical office measures. These inflation rates always cover the period from March in the previous year to March in the current year. Most annual inflation rates are close to 0% with somewhat more positive rates.

Figure D1: Absolute Frequencies of Annual Inflation Rates per Product Class (March 2010 – March 2015)[1]
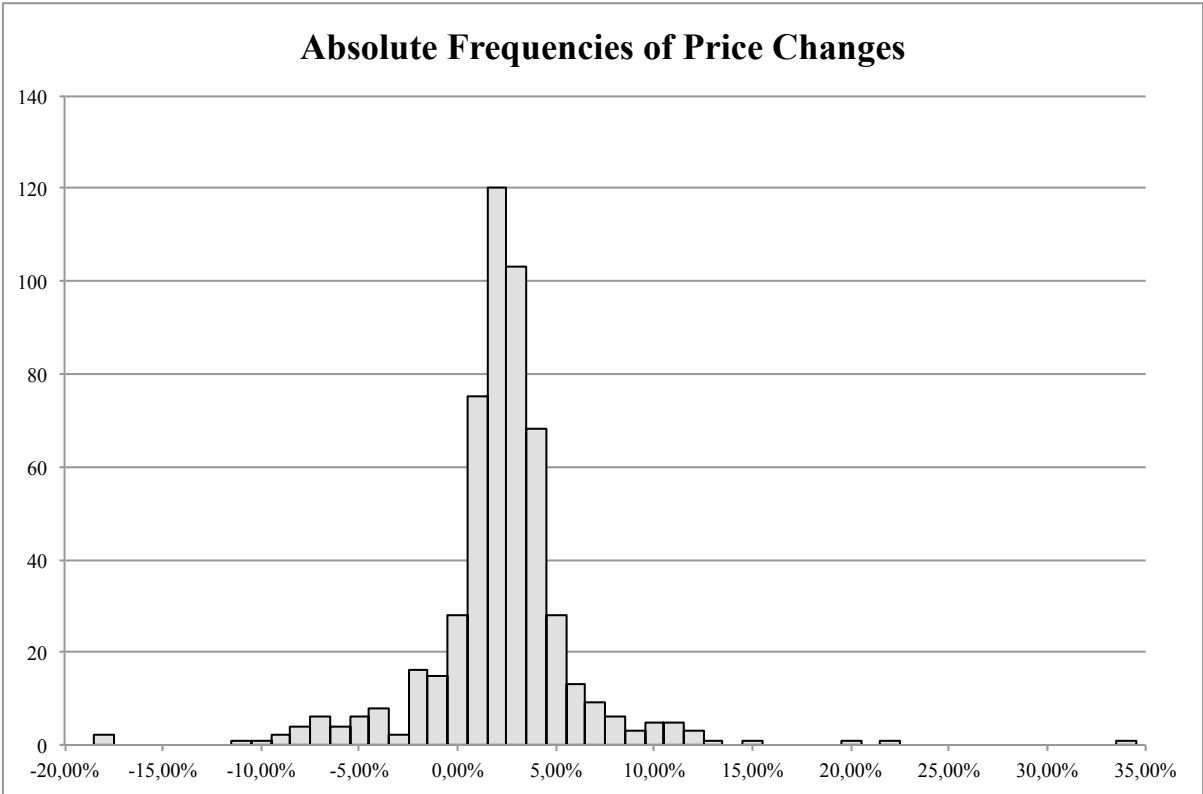
1) values on the x-axis depict class midpoints.

Appendix E: Overview of Symbols

Table E1: Overview of Symbols

Variables	
AOS	Affinity for online shopping
BL	Brand loyalty
CC	Category clutter
CE	Customer expertise
DPC	Direction of price change
DUR	Dummy variable indicating that a product class contains mainly durables
ES	Economic skills
FRE	Purchase frequency
MI	Monetary impact
MS	Mathematical skills
NI	Net income
NPC	Number of product classes
PC	Price consciousness
PCR	Price change range
PI	Perceived inflation (scale value)
PPPF	Perceived price promotion frequency
PQR	Price-quality relationship
REC	Purchase recency
SA	Level of social amplification
SER	Dummy variable indicating that a product class contains mainly services
SEX	Dummy variable indicating that a person is male
Sociodemographics	Vector containing the gender (male=1), age, income, income change (decrease=-1, stable=0, increase=1), individuals' involvement in daily shopping activities, household size, education (graduate=1), profession (student=1), and martial status (living in a partnership=1).
TSQ	Time spent for the questionnaire
π	Inflation rate (percentage value)
π^{p}	Perceived inflation rate (percentage value)
Indexes	
i	Product group index with $i = 1,\ldots I$ (number of product groups)
j	Product class index with $j = 1,\ldots J$ (number of product classes in a product group)
k	Respondent index with $k=1,\ldots K$ (number of respondents)
s	Stimulus index with $s=1,\ldots S$ (number of stimuli)
Model parameters	
α, β, γ	Regression parameters to be estimated
ε, ζ, v	Error terms
Σ	Variance-covariance matrix of random parameters
w	Weighting parameter for weighted models
Other Abbreviations	
COICOP	Classification of individual consumption according to purpose

Appendix F: Model Validation for the Mapping Function

We compare the forecast performance of our mapping function (equations (1) – (3)) against several other models. We randomly split our sample into two equal subsamples, an estimation sample and a holdout sample. We estimate model variants as below on the estimation sample and check their forecast performance.

Other model variants are as follows: Model T1 is a standard linear model without any heterogeneity. It tests whether the transformation function is individual. Model T2 involves heterogeneity by a random intercept and a random coefficient. It checks whether respondents' sociodemographics reveal explanation. In model T3, we add a reduced set of interactions with sociodemographics. It tests whether our choice of sociodemographics is meaningful. Model T4 contains the full set of demographics, but only involves heterogeneity and interactions in the coefficient β_k. It checks whether individual characteristics' only intervene at the coefficient.

Table F1 shows an overview on the models and their forecast performance. Our mapping function (equations (1) – (3)) leads to the lowest Mean Absolute Error and Mean Squared Error for both, the estimation and holdout sample.

able F1: Comparison of Models for the Mapping Function

| Modell | Modell Components | | | | Forecast Performance | | | |
| | Intercept | | Stimulus | | Estimation Sample | | Holdout Sample | |
	Random	Interaction Sociodemographics	Random	Interactions Sociodemographics	MAE	MSE	MAE	MSE
Mapping Function	✓	✓	✓	✓	2.053	6.987	2.040	6.909
T1	-	-	-	-	2.110	7.119	2.071	6.985
T2	✓	-	✓	-	2.093	7.147	2.057	7.006
T3	✓	selective[1]	✓	selective[1]	2.072	7.076	2.049	6.972
T4	-	-	✓	✓	2.067	7.121	2.048	7.044

1) only interactions with gender, age, income, profession, and education included.

Appendix G: Estimation of the Perceived Market Basket Composition

Table G1 below provides the estimation results for the perceived market basket composi-tions. We estimate these compositions according to equations (5) – (7) drawing from our main article. Despite the general approach, as table W1 above shows, product groups consist of different numbers of product classes. For example, the product group personal care divides into three product classes, while the group transportation services consists of six product clas-ses. Therefore, the number of independent variables in equation (5) of the main article and, thus, the number of parameters to estimate differ. Correspondingly, the number of estimated parameters in table W5 below varies across product groups.

Table G1: Estimation Results for the Perceived Market Basket Composition

Product class[1)]	Coefficient	Product group					
		Alcoholic beverages		Audio-visual, photo-graphic, and infor-mation processing equipment		Clothing	
Intercept	α_i	-.052	(.154)	.030	(.239)	.224 **	(.106)
	SD	1.327 ***	(.145)	.333	(.213)	.894 ***	(.096)
1[1]	$\ln(\beta_{i1})$	-.558 **	(.218)	-.848	(.578)	-.982 ***	(.252)
	ln(variance)	.012	(.187)	.028	(.507)	.039	(.206)
2[1]	$\ln(\beta_{i2})$	-1.665 ***	(.606)	-2.253	(2.477)	-1.402 ***	(.330)
	ln(variance)	.098	(.350)	.009	(1.517)	.220	(.149)
3[1]	$\ln(\beta_{i3})$	-4.078	(5.352)	-2.670 *	(.979)	-2.670 ***	(.919)
	ln(variance)	.007	(3.859)	.245	(.432)	.077	(.448)
4[1]	$\ln(\beta_{i4})$	-2.919	(1.992)	-2.977 **	(1.514)	-1.940 ***	(.358)
	ln(variance)	.012	(.711)	.014	(.999)	.049	(.213)
5[1]	$\ln(\beta_{i5})$	-		-2.827	(2.175)	-	
	ln(variance)			.036	(1.551)		
N		101		96		102	
Log Likelihood		-208.61666		-175.67211		-178.94944	
Pseudo-R^2		.895		.639		.942	

Table G1: Estimation Results for the Perceived Market Basket Composition (Cont'd)

Product class[1]	Coefficient	Product group		
		Electricity, gas, and other fuels	*Household appliances*	*Insurance services*
Intercept	α_i	.057 (.194)	-.257 * (.148)	-.117 (.156)
	SD	.301 (.185)	.484 *** (.153)	.067 (.155)
1[1]	$\ln(\beta_{i1})$	-.528 *** (.186)	-.927 *** (.329)	-.837 *** (.272)
	ln(variance)	.089 (.158)	.297 (.211)	.232 (.214)
2[1]	$\ln(\beta_{i2})$	-2.910 (2.672)	-1.109 *** (.419)	-2.553 * (1.399)
	ln(variance)	.143 (1.573)	.208 (.193)	.066 (.745)
3[1]	$\ln(\beta_{i3})$	-1.578 *** (.599)	-3.189 *** (.697)	-3.010 *** (.878)
	ln(variance)	.231 (.251)	.006 (.418)	.100 (.664)
4[1]	$\ln(\beta_{i4})$	-2.456 *** (.621)	-	-1.440 *** (.215)
	ln(variance)	.012 (.566)		.106 (.149)
5[1]	$\ln(\beta_{i5})$	-2.999 (2.888)	-	-
	ln(variance)	.003 (1.554)		
N		97	101	96
Log Likelihood		-175.72987	-185.39395	-173.12786
Pseudo-R^2		.810	.738	.732

Product class[1]	Coefficient	Product group		
		Medical products, appliances, and equipment	*Newspapers, books, and stationery*	*Operation of personal transport equipment*
Intercept	α_i	-.057 (.114)	-.277 (.173)	-.103 (.167)
	SD	1.000 *** (.104)	.476 *** (.165)	.769 *** (.139)
1[1]	$\ln(\beta_{i1})$	-.594 *** (.229)	-.862 *** (.199)	-1.018 *** (.338)
	ln(variance)	.307 * (.179)	.115 (.162)	.025 (.296)
2[1]	$\ln(\beta_{i2})$	-1.952 *** (.669)	-2.022 ** (.787)	-2.636 ** (1.073)
	ln(variance)	.110 (.368)	.182 (.424)	.155 (.448)
3[1]	$\ln(\beta_{i3})$	-2.389 ** (1.205)	-3.421 (3.504)	-1.540 *** (.580)
	ln(variance)	.033 (.388)	.046 (2.286)	.034 (.402)
4[1]	$\ln(\beta_{i4})$	-	-1.781 *** (.643)	-1.320 *** (.445)
	ln(variance)		.006 (.396)	.022 (.275)
N		102	100	94
Log Likelihood		-187.64332	-195.53710	-167.91274
Pseudo-R^2		.918	.650	.792

Table G1: Estimation Results for the Perceived Market Basket Composition (Cont'd)

Product class[1]	Coefficient	Product group		Personal care		Recreational and cultural services	
		Other recreational items and equipment, gardens and pets					
Intercept	α_i	-.201	(.230)	-.169 ***	(.053)	-.229	(.181)
	SD	.062	(.210)	1.632 ***	(.050)	.364 **	(.183)
1[1]	$\ln(\beta_{i1})$	-1.848 ***	(.707)	-.603 ***	(.080)	-.320 **	(.145)
	ln(variance)	.153	(.571)	.116 **	(.057)	.035	(.132)
2[1]	$\ln(\beta_{i2})$	-1.816 **	(.920)	-1.722 ***	(.271)	-2.517 *	(1.365)
	ln(variance)	.040	(.685)	.084	(.116)	.042	(.967)
3[1]	$\ln(\beta_{i3})$	-.598 ***	(.223)	-1.352 ***	(.185)	-1.838 ***	(.447)
	ln(variance)	.028	(.211)	.183 **	(.083)	.042	(.334)
4[1]	$\ln(\beta_{i4})$	-2.892	(1.824)	-		-	
	ln(variance)	.001	(1.224)				
5[1]	$\ln(\beta_{i5})$	-2.867	(2.122)	-		-	
	ln(variance)	.019	(1.461)				
N		97		92		99	
Log Likelihood		-186.23143		-176.13170		-181.46064	
Pseudo-R^2		.643		.998		.691	

Product class[1]	Coefficient	Product group	
		Transportation services	
Intercept	α_i	.096	(.227)
	SD	.137	(.211)
1[1]	$\ln(\beta_{i1})$	-1.228 **	(.583)
	ln(variance)	.376	(.419)
2[1]	$\ln(\beta_{i2})$	-2.399 *	(1.273)
	ln(variance)	.093	(.948)
3[1]	$\ln(\beta_{i3})$	-2.668	(1.733)
	ln(variance)	.004	(1.451)
4[1]	$\ln(\beta_{i4})$	-2.980	(2.267)
	ln(variance)	.029	(1.461)
5[1]	$\ln(\beta_{i5})$	-1.505 *	(.776)
	ln(variance)	.227	(.321)
6[1]	$\ln(\beta_{i6})$	-3.063	(2.262)
	ln(variance)	.046	(1.240)
N		104	
Log Likelihood		-215.80899	
Pseudo-R^2		.557	

1) product class depicts the number of product class per product group as given in table W1 in appendix A.

Appendix H: Descriptives of the Results

Table H1 contains the descriptives of our driver analysis. It provides the minimum, maximum, mean value, and standard deviation for the potential drivers.

Table H1: Descriptives for Product Class Averages

	N	Min	Max	Mean	SD
Deviation Perceived-Official Inflation	3,930	-.795	.960	.070	.328
Average Weight[1]	4,996	.000	4.192	.942	.723
Product class Characteristics					
Price Change Range	52^2	.000	.350	.041	.066
Durables	53	.000	1.000	.283	.455
Services	53	.000	1.000	.434	.500
Direction of Price Change	52^2	-1.000	1.000	.442	.873
Size of Price Change	52^2	.000	.191	.030	.039
Category Clutter	53	1.000	35.000	4.830	5.843
Official Inflation	52^2	-.191	.188	.009	.048
Individual Characteristics in a Product Class					
Purchase Frequency	4,996	.020	3.000	.222	.509
Purchase Recency	4,996	-730.000	-1.500	258.830	292.208
Perceived Price Promotion Frequency	4,996	1.000	7.000	3.770	1.790
Price Consciousness	4,996	1.000	7.000	4.269	1.950
Affinity for Online Shopping	4,996	1.000	7.000	2.952	2.020
Brand Loyalty	4,996	1.000	7.000	3.636	1.861
Customer Satisfaction	4,996	1.000	7.000	4.240	1.559
Customer Expertise	4,996	1.000	7.000	3.633	1.797
Price-Quality Relationship	4,996	1.000	7.000	2.975	1.567
Social Amplification	4,996	1.000	7.000	2.779	1.626
Monetary Impact in €	4,996	5.000	750.000	105.339	172.130

1) adjusted by multiplication with the number of product classes of a product group; 2) for the product class mixed drinks (<6% alcoholic strength) no official inflation values were available.

REFERENCES APPENDIX PROJECT 1

Ailawadi, Kusum L., Scott A. Neslin, and Karen Gedenk (2001). "Pursuing the Value-conscious Consumer: Store Brands versus National Brand Promotions," *Journal of Marketing*, 65 (1), 71-89.

Burton, Scot, Judith A. Garretson, and Anne M. Veliquette (2012). "Implications of Accurate Usage of Nutrition Facts Panel Information for Food Product Evaluations and Purchase Intentions," *Journal of the Academy of Marketing Science*, 27 (4), 470-480.

Chang, C. (2004). "The Interplay of Product Class Knowledge and Trial Experience," *Journal of Advertising*, 33 (1), 83-92.

Darden, William R. and William D. Perreault (1976). "Identifying Interurban Shoppers: Multiproduct Purchase Patterns and Segmentation Profiles," *Journal of Marketing Research*, 13 (1), 51-60.

Drake, Peter R. and Darren W. Dahl (2003). "Fairness and Discounts: The Subjective Value of a Bargain," *Journal of Consumer Psychology*, 13 (3), 328-338.

Estelami, Hooman and Peter de Maeyer (2004). "Product Category Determinants of Price Knowledge for Durable Goods," *Journal of Retailing*, 80 (2), 129-137.

EU Consumer Survey (2016). *Questionnaires in National Languages* [available at http://ec.europa.eu/economy_finance/db_indicators/surveys/questionnaires/index_en.htm; accessed December 13, 2015].

Gaski, John F. and Michael J. Etzel (1986). "Index of Consumer Sentiment Toward Marketing," *Journal of Marketing*, 50 (3), 71-81.

Homburg, Christian, Nicole Koschate- Fischer, and Christian Wiegner (2012). "Customer Satisfaction and Elapsed Time since Purchase as Drivers of Price Knowledge," *Psychology and Marketing*, 29 (2), 76-86.

Kozup, John C., Elizabeth H. Creyer, and Scot Burton (2003). "Making Healthful Food Choices: The Influence of Health Claims and Nutrition Information on Consumers' Evaluations of Packaged Food Products and Restaurant Menu Items," *Journal of Marketing*, 67 (2), 19-34.

Lundstrom, William J. and Lawrence M. Lamont (1976). "The Development of a Scale to Measure Consumer Discontent," *Journal of Marketing Research*, 13 (4), 373-81.

Monroe, Kent B. (2003), Pricing: Making Profitable Decisions, Burr Ridge: McGraw-Hill.

Strutton, David and Jeffrey Lewin (2012). "Investigating Consumers' Responses to the Great Recession," *Journal of Consumer Marketing*, 29 (5), 378-388.

United Nations Statistics Division (2016). *Classification of Individual Consumption According to Purpose* [available at http://unstats.un.org/unsd/cr/registry/regcst.asp?Cl=5; accessed at November].

Yamamuchi, Kent T., and Donald I. Templer (1982). "The Development of a Money Attitude Scale," *Journal of Personality Assessment*, 46 (5), 522-528.

PROJECT 2: DRIVERS OF B2B BRAND STRENGTH –

INSIGHTS FROM AN INTERNATIONAL STUDY ACROSS INDUSTRIES

Author: Eric Michael Lennartz, Marc Fischer, Manfred Krafft, and Kay Peters

Acknowledgment

ABSTRACT

This study analyzes the effect of brand associations and marketing-mix instrument perceptions on brand strength for B2B firms. Although B2B brands may contribute substantially to firm profit, only little research has been directed at them. We close this research gap by analyzing a unique data set that spans across three countries and seven industries. We find that the brand-associations 'sustainability and corporate governance' as well as 'innovation and expertise' drive brand strength in B2B markets across all countries and industries. For marketing-mix instruments, product and distribution perceptions shape brand strength. However, the effects of marketing-mix instrument perceptions vary by industry and country.

Keywords: Business-to-Business Markets; Brand Strength; Brand Associations; Country Effects; Industry Effects.

1 Introduction

Studies on brand research predominantly focus on business-to-consumer (B2C) branding (e.g., Homburg, Klarmann, and Schmitt 2010; Sethuraman, Tellis, and Briesch 2011). However, according to Interbrand (2013) four of the top ten global brands — GE, IBM, Intel, and Microsoft — are primarily business-to-business (B2B) brands. A recent meta-analysis of 3,800 brands by Fischer and Himme (2013) underlines their importance: the average B2B brand has a median value of US$1,135bn compared to US$1,595bn for B2C brands. In addition, since the 1990s, the growth path appears to be less volatile for B2B brands than for B2C brands, thus increasing their relative value compared to B2C brands. These findings suggest that branding may be a powerful option to create economic value and a competitive advantage in B2B markets as well.

In a recent contribution, Wiersema (2012) supports the fact that the importance of branding has increased in B2B markets. Based on in-depth interviews of 72 executives and 30 leading B2B researchers, Wiersema identifies brand management, sales management, and customer centricity as the major topics in current B2B marketing. However, in contrast to its importance, Wiersema identifies B2B brand management as the most under-researched and thus the most promising area for research. Our own review on the B2B brand management literature supports his finding. We add that identifying drivers of brand strength and assessing their relative impact across countries and industries represents one of the most prominent gaps in the B2B brand management literature.

From the practitioner's point of view, industrial suppliers identify brand management as a promising area that can be used to improve firm performance (e.g., Gosh and John (2009)). A general question is whether there are any substantial differences between B2C and B2B brand management. Without such differences, managers could simply leverage B2C branding insights to improve B2B brand management. But an analysis of the B2B brand management

literature review reveals that, while the concept of perceived brand strength may be applied in both domains, there are indeed substantial differences with respect to drivers of brand strength requiring a distinct B2B branding approach.

Incorporating these distinctions, we first develop a holistic as well as parsimonious conceptual model of B2B brand strength and its drivers that generalizes across both, countries and industries. Our model also introduces two important dimensions of brand associations that have been neglected previously. Second, we test this new model in a large-scale survey across three countries and seven industries. Such an international and cross-industry perspective represents a unique research opportunity and constitutes a key feature of our empirical analysis. Finally, we derive important implications from our findings for both, researchers and practitioners.

The paper proceeds as follows: In section 2, we develop a conceptual model for drivers of B2B brand strength, followed by the empirical analysis in section 3. Section 4 covers our estimation and findings. Section 5 outlines the implications for research and management before ending on the concluding remarks in section 6.

2   Conceptual Background

2.1 Prior Research

Aaker and Jacobson (2001) show that brand attitudes drive stock returns and brand strength in high-technology B2B markets. Homburg, Klarmann, and Schmitt (2010) report effects of brand awareness on market performance in a cross-sectional B2B setting. In a qualitative approach, Fischer, Giehl, and Freundt (2011) identify potential brand attributes associated with the DHL brand that influence customers' evaluations of that B2B brand. All these studies indicate that the general concept of brand value chains also holds in various selected domestic B2B contexts. However, we lack a generalizing evidence across industries and

countries that explicitly links brand associations to brand strength in B2B markets. For example, does brand strength result from intangible value-oriented dimensions (such as brand associations) or rather from performance dimensions (such as product perceptions by customers). Accordingly, leveraging such a broad cross-country and cross-industry sample could contribute substantially to our understanding of the relationship between brand associations and brand strength in B2B markets.

When investigating potential brand associations, Brown and Dacin (1997) show that corporate associations should be initially separated into corporate social responsibility (CSR) and corporate ability. In our context, the former might entail 'sustainability and corporate governance' (e.g., Torres et al. 2012), especially if a firm acts as a supplier. Due to potential ethical concerns of followers within a supply chain, a focus on low prices should not come at the expense of sustainability and corporate governance. In a B2B context, and in line with Aaker and Jacobson (2001), the latter aspect of corporate ability might result from 'innovation and expertise', since high risks inherent to B2B buying decisions might be lowered by such capabilities. But how should these two B2B brand associations be measured? Previous studies investigate brand associations only in qualitative settings (e.g., Fischer, Giehl, and Freundt 2011). As a consequence, we need to develop a new approach to assess such latent structures of brand associations for our quantitative study. This is essential to later derive managerial guidelines concerning desirable brand profiles. Accordingly, we assess the underlying structure of brand associations in B2B markets through survey data.

Although a few studies report on effects of marketing-mix instrument perceptions on B2B brand strength (e.g., Ghosh and John 2009), generalizable evidence is again missing. In particular, there is a knowledge gap concerning the simultaneous impact of individual marketing-mix instrument perceptions. Homburg, Klarmann, and Schmitt (2010), and Fischer, Giehl, and Freundt (2011) cover only selected marketing-mix instruments. However, e.g., potential

rationality in buying center structures might favor product perceptions over that of communication activities. Hence, it is important to uncover which marketing mix instrument perceptions actually affect B2B brand strength and how they perform in relation to each other. The unique and dedicated survey underlying this study allows studying both, the effect of brand associations and all marketing-mix perceptions on brand strength, within a single framework.

B2B brands are used in numerous industries and are subject to global market structures (Torres et al. 2012). Previous research covers only selected industry or country effects, which again limits the generalizability of such results. Building on that research, our study accounts for various country- and industry-specific covariates that may frame the performance of drivers on brand strength.

2.2 Conceptual Model of Drivers of B2B Brand Strength

The American Marketing Association (2013) defines "…a brand [as] a customer experience represented by a collection of images and ideas; often, it refers to a symbol such as a name, logo, slogan, and design scheme. Brand recognition and other reactions are created by the accumulation of experiences with the specific product or service, both directly relating to its use, and through the influence of advertising, design, and media commentary". This definition incorporates both B2C and B2B businesses. If we transfer this to B2B markets, we can define a brand as a symbol "…considered by industrial players as an indispensable reference in conjunction with a particular need" (Kapferer 2008, p. 113).

A key element of brand investments that translates into growth and higher profits is the perceived strength of a brand, i.e., its familiarity, knowledge, and customers' perception of its performance (Hoeffler and Keller 2003); stronger brands may help firms to lower price sensitivity (e.g., Erdem, Swait, and Louviere 2002; Kalra and Goodstein 1998), establish brand extensions (e.g., Rühle et al. 2012), or improve consideration in customers' purchase processes (e.g., Lehmann and Pan 1994). Although researchers find these results mainly for B2C

markets, initial evidence suggests that brand strength might also generate competitive advantages in B2B markets (e.g., Ghosh and John 2009). Thus, brand strength seems to represent an adequate expression of brands' success. Therefore, in our framework we propose brand strength as our dependent variable.

Consistent with the idea of a brand value chain (Keller and Lehmann 2006), we assume that brand associations are an antecedent of brands' success. Brand associations combine attributes, benefits, and attitudes related to a brand. In B2C markets, researchers can, for example, chose among dimensions such as excitement, sincerity, sophistication, competence, and ruggedness (Monga and Lau-Gesk 2007). These characteristics only partly fit into B2B markets. Moreover, research indicates that the associated abilities of a firm and customers' associations concerning a firm's sense of responsible acting are critical in B2B markets (Brown and Dacin 1997). Thus, the items for the brand association scale that we report in Table 1 reflect ability as well as responsibility concerns.

In line with recent studies on B2C branding (e.g., Stahl et al. 2012), we include all marketing-mix instrument perceptions in our framework. The general structure of the marketing mix, i.e., product, distribution, price, and communication, should not differ between B2B and B2C markets. To assure that marketing mix perceptions are not redundant antecedents of brand strength, we define all items for the marketing instrument scales without any relation to brand perception (e.g., Ataman, van Heerde, and Mela 2010; Keller and Lehmann 2006).

Homburg, Klarmann, and Schmitt (2010) report that the market performance of brands may be framed by both, industry and firm characteristics. Hence, we include both elements as controls in our framework. When researchers analyze brands instead of strategic business units, firm characteristics may differ to some extent. Past research indicates that the researcher needs to control for firm size and role of the respondent within the firm (Homburg et al. 2012; Zablah, Brown, and Donthu 2010). Respondents' brand strength perceptions might also

differ according to their involvement within a firm's purchase process (Keller and Lehmann 2006). In addition, most firms in B2B markets act globally, and so there might be variation between countries (Torres et al. 2012). Hence, we include both the country and the international business reach of firms in our framework. Figure 1 depicts the resulting framework.

Table 1: Items of the Brand Strength, Marketing Mix, and Brand Associations

Brand Strength	*Product*
1. A company with good reputation	1. Wide variety of different products or services (breadth of product/service offer)
2. A company I trust	2. A lot of meaningful value added to the products or services (e.g., warranties, support)
3. A company I feel positive about	3. Exactly the products/services that meet to my needs
4. A company that I know well	4. Ahead of the market through constant product innovation
5. A company that is clear on what it stands for	5. Good quality products/services
Brand Association Items	*Distribution*
1. Cares about an honest and open dialog with its customers, investors, and society	1. Convenient access to the company's offering whenever I need it
2. Role models corporate social responsibility	2. Fast response times
3. Promotes and practices sustainability in its products or services	3. Sales representatives and personnel with the necessary expertise and knowledge
4. Promotes diversity and equal opportunities	4. Strong relationship with the dealer/supplier of this company
5. Acts responsibly across its supply chain	5. Timely and reliable delivery of products/services
6. Fits in well with my values and beliefs in its work	*Price*
7. Is a driver of innovation	1. Lower prices than comparable products or services from other manufacturers
8. Has a high level of expertise	2. Good value for money
9. Shapes the direction of the market	3. Fair and transparent pricing policies and price adjustments over time
10. Is a leader in its field	4. Cheapest cost for product maintenance
11. Has global reach	5. Attractive terms and conditions (e.g., discounts, financing conditions)
	Communication
	1. Provides the information I need
	2. Keeps me informed about relevant new topics
	3. Highly visible and present in media
	4. A relevant source of industry-specific information
	5. Offers a wide range of possibilities for online interaction

Figure 1: Conceptual Framework for Brand Strength in B2B Markets
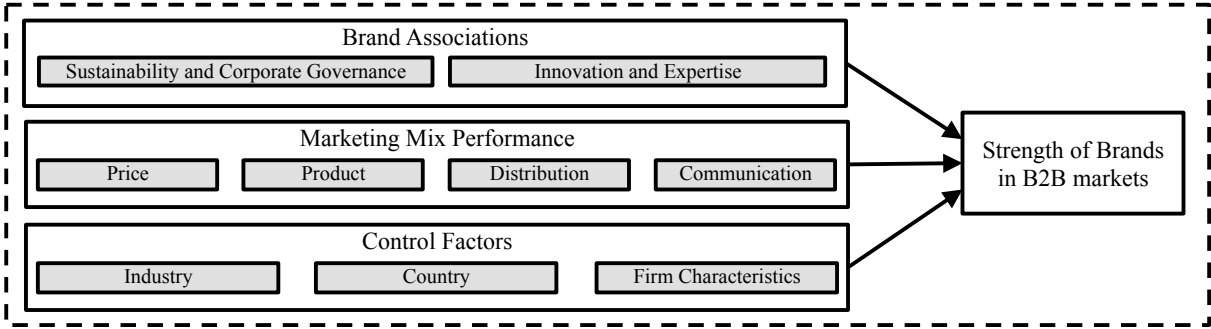
3   Empirical Analyses

3.1 Measurements of Constructs

We obtain our data by using an international online survey among managers. For all constructs, we operationalize, validate, and measure scales by following the approach suggested by Churchill (1979).

We use multi-item scales to measure brand strength, brand associations, and performance perceptions of the four marketing-mix instruments, price, product, distribution, and communication. For all scales, we generate item pools by screening the relevant literature (e.g., Fischer 2007) and by conducting qualitative interviews. We measure all items via fully labeled five-point Likert scales.

We pretest all items for all constructs in focus groups with experts in B2B branding. We adapt a few items from their evaluation of the items' relevance, clarity, and specificity.

Brand strength pretests indicate that a few items differ in their importance because of firm, industry, and country specifics. Hence, we asked respondents to select the three items out of the five-item battery that were most important to them. This procedure leads to eight potential combinations of three items, which differ in their relative frequencies. Accordingly, we check whether all eight potential item combinations reflect the same construct brand

strength. We do so by computing Cronbach's alpha for each of the eight present item combinations in the sample. As a rule, we require that the average Cronbach's alpha across all combinations should exceed 0.8. In addition, it should be above 0.7 for at least 80% of all eight item-combinations. These criteria are in line with recommendations in the relevant literature (e.g., Peterson 1994). All average values of reliabilities for the brand strength and the four marketing-mix scales are above 0.8 and more than 80% of the potential item selections reveal alpha values beyond 0.7 for the scales to be checked. Table 2 gives the details. As these criteria are fulfilled, we use the average values of the three items selected by the respondents.[2]

Table 2: Reliability Criteria Based on Coefficient Alpha for the Different Item Combinations of the Scales

Criterion	Brand strength	Distribution	Price	Product	Communication
Weighted average α	0.82	0.82	0.82	0.80	0.84
Weighted proportion of item combinations with $\alpha > .7$	100%	100%	96%	94%	100%

Note: For each scale, respondents chose 3 out of 5 available items yielding 8 possible item combinations. For all combinations, we computed coefficient alpha. Weighted average alpha is the average coefficient alpha across all possible combinations. The weight measure the relative frequency a 3-item combination was chosen among respondents.

For the brand association constructs, we apply an exploratory factor analysis to the 11 items listed in table 1 to test for their underlying structure. We apply the scree criterion and select a two-factor solution accounting for 62.6% of the total variance of the items (see table 3). According to the item loadings, the two factors reflect the latent brand association structure within our framework. Factor 1 consists primarily of items representing 'sustainability and corporate governance'. Items that load on Factor 2 reflect 'innovation and expertise'. The resulting coefficient alphas are 0.884 for the items loading on Factor 1 and 0.835 for the items

---

[2] Respondents choosing three out of five items that are most relevant to them is not implying that they perceive different constructs. We pre-test this extensively. Responding to fewer items reduces respondents' burden. Since the underlying construct is reflective it does not matter if items are deleted.

loading on Factor 2. The core brand associations appear to be represented by these two factors. Hence, we use standardized factor scores for the rotated factor solution.

Table 3: Factor Loadings of the Exploratory Factor Analysis

Item	Factor 1 Sustainability and Corporate Governance	Factor 2 Innovation and Expertise
Cares about an honest and open dialog with its customers, inventors, and society	**0.766**	0.264
Role models corporate social responsibility in its work	**0.759**	0.282
Promotes and practices sustainability in its products or services	**0.746**	0.286
Promotes diversity and equal opportunities	**0.723**	0.301
Acts responsibly across its supply chain	**0.715**	0.295
Fits in well in my values and believes	**0.715**	0.309
Is a driver of innovation	0.471	**0.608**
Has a high level of specialist expertise	0.435	**0.612**
Shapes the direction of the market	0.381	**0.664**
Is a leader in its field	0.385	**0.716**
Has global reach	0.106	**0.838**

Both factors explain 62.6% of total variance

Extraction method: principal component analysis; rotation method: Varimax with Kaiser normalization.

Pretests on the four marketing-mix instrument perceptions again show that a few items differ in their importance due to firm, industry, and country specifics. Hence, we follow the same procedure as for brand strength. Again, reliabilities for the four marketing-mix instrument perceptions are above 0.8 and more than 80% of the potential item selections reveal alpha values beyond 0.7 (see table 2). As all required criteria are fulfilled, we use the average values of the three items selected by the respondents for the specific marketing-mix instrument perception for each respondent.

We operationalize the controls as follows (see table A1 in appendix A for details). We categorize respondents by their role in the decision process, ranging from being a final deci-

sion maker to being only part of a recommendation team. Business reach describes the geographic coverage of the respondent's firm, i.e., whether it is a regionally or globally active firm. The share of revenues describes how much of a firm's sales is actually generated in B2B markets, i.e., from exclusively B2B to exclusively B2C, representing an intermediate B2B firm or one that deals with end consumers. We assess firm size in five categories (100-249, 250-999, 1,000-4,999, 5,000-9,999, and more than 10,000 employees). We transform this scale into a metric scale representing relative differences between employee categories.

3.2 Data sample and description

We investigate seven industries across three countries. Table 4 provides details on the industry and country structure; we study 'finance and insurance' and 'utilities' only for Germany and the US. We choose the three countries, Germany, the US, and India, to cover an adequate spectrum of cultural, economic, and political diversity. These three countries are especially relevant since they cover about 30% of the World's total GDP in 2012 (World Bank 2013). The investigated industries represent the most relevant B2B sectors throughout these economies, contributing more than 40% of the US GDP in 2011 (Bureau of Economic Analysis 2013). We choose this combined diversified setup to ensure the potential generalizability of results across countries and industries. It also allows us to quantify industry- and country-specific effects.

We collect the data by using an online survey performed by a global market research firm during August and September of 2012. We choose an online survey since it is particularly appropriate in cross-national research (e.g., Fischer, Völckner, and Sattler 2010). Respondents were decision makers in B2B firms. To secure the relevance of B2B branding related topics, we require these firms to have at least 100 employees.

Each respondent evaluated, once for a primary and once for a secondary provider, all items for brand strength, brand associations, and the marketing-mix performance within the

industry considered. We ensured that respondents referred to providers they had already dealt with. For example, we checked for respondents' brand knowledge by asking for providers' brand names. Integrating respondents' primary and secondary provider has several advantages. By doing so we expand our sample and add variability on the preference of brands. As described above, respondents had to pre-select the most important three items for brand strength and each marketing-mix instrument before answering. Thus, the sample included 2,180 observations from 1,090 respondents.

We eliminate 144 subjects with 288 cases (13.2% of 1,090) for which either major parts of the information are missing, response style is inadequate, or the respondent had spent less than five minutes on the questionnaire (with an average response time of 15 minutes). We define an inadequate response style as one that includes scale mix-ups as well as click-through behavior. Table 4 reports the sample structure of the remaining 1,892 observations by country, industry, and firm size.

Table 4: Sample Structure by Country, Industry, and Firm Size

Country (N)	Industry (N)	Firm size in employees (N)
Germany (660); India (652); USA (580)	Chemicals, commodities, and basic materials (290); IT (322); communication (320); logistics (300); machinery and equipment (302); finance and insurance (174); utilities (184)	100-999 (428); 1000-9999 (922); ≥10000 (528)
Σ 1,892 obs.	Σ 1,892 obs.	Σ 1,878 obs.

Table 5 shows the means and standard deviations on a general level, by country, and by industry. We find it interesting that the mean of brand strength (range 3.672-3.982) does not vary much across countries and industries, and that all values are within each other's standard deviations. Since we cover mostly larger firms (i.e., those with more than 100 employees) those firms should tend to have suppliers that compete in the international market place, so that the similarities may be a reflection of globalization. However, the variation is substantial: the standard deviations are in the range of 17% to 25%.

Table 5: Descriptive Statistics of Constructs

Variable	Overall	Germany	India	USA	Chemicals, Commodities, and basic materials	IT	Telecommu- nications	Logistics	Machinery and equip- ment	Finance and Insurance	Utilities
Brand Strength	3.861 (0.758)	3.805 (0.745)	3.982 (0.709)	3.789 (0.810)	3.886 (0.718)	3.951 (0.679)	3.819 (0.822)	3.858 (0.729)	3.974 (0.779)	3.741 (0.728)	3.672 (0.829)
Sustainability and corporate govern- ance	3.593 (0.692)	3.340 (0.671)	3.835 (0.675)	3.606 (0.631)	3.670 (0.653)	3.606 (0.647)	3.482 (0.722)	3.680 (0.673)	3.758 (0.689)	3.436 (0.644)	3.377 (0.748)
Innovation and expertise	3.715 (0.686)	3.594 (0.689)	3.850 (0.683)	3.702 (0.659)	3.791 (0.623)	3.791 (0.682)	3.609 (0.721)	3.780 (0.675)	3.892 (0.646)	3.548 (0.675)	3.208 (0.669)
Product performance	3.726 (0.747)	3.635 (0.734)	3.870 (0.726)	3.668 (0.761)	3.789 (0.727)	3.797 (0.724)	3.699 (0.764)	3.736 (0.706)	3.821 (0.796)	3.598 (0.710)	3.502 (0.748)
Distribution performance	3.750 (0.765)	3.640 (0.769)	3.900 (0.720)	3.700 (0.785)	3.780 (0.744)	3.800 (0.744)	3.620 (0.800)	3.810 (0.747)	3.850 (0.741)	3.630 (0.776)	3.650 (0.793)
Price performance	3.617 (0.773)	3.435 (0.734)	3.858 (0.731)	3.553 (0.793)	3.706 (0.700)	3.653 (0.778)	3.539 (0.819)	3.710 (0.742)	3.711 (0.752)	3.423 (0.815)	3.431 (0.765)
Communication performance	3.640 (0.762)	3.480 (0.760)	3.800 (0.741)	3.660 (0.750)	3.680 (0.795)	3.630 (0.742)	3.590 (0.787)	3.720 (0.745)	3.710 (0.761)	3.600 (0.724)	3.490 (0.742)
Role of respondent	.460 (0.499)	.415 (0.493)	.607 (0.489)	.345 (0.476)	.448 (0.498)	.385 (0.487)	.438 (0.497)	.547 (0.499)	.470 (0.500)	.448 (0.499)	.500 (0.501)
Business reach	1.700 (0.940)	1.620 (0.880)	1.700 (0.891)	1.810 (1.051)	1.620 (0.901)	1.570 (0.905)	1.820 (1.002)	1.730 (0.921)	1.600 (0.929)	1.840 (0.906)	1.870 (0.976)
Revenue from B2B compared to B2C	2.640 (1.191)	2.450 (1.201)	2.530 (1.009)	2.970 (1.316)	2.370 (1.093)	2.480 (1.249)	2.830 (1.216)	2.590 (1.100)	2.630 (1.157)	2.880 (1.201)	2.880 (1.276)
Firm Size	30.644 (21.383)	33.187 (22.263)	23.842 (17.791)	35.370 (22.163)	28.236 (19.472)	34.340 (21.210)	32.750 (22.650)	30.880 (21.373)	30.550 (20.931)	26.709 (20.472)	27.780 (22.690)

Standard deviations in parentheses; Mean values and standard deviations for sustainability and corporate governance as well as for innovation and expertise are computed on the items dominantly loading on the factor according to the separation in table 3.

In comparison, the brand associations show a larger variation across countries and industries. For sustainability and corporate governance, the means range from 3.34 in Germany to 3.835 in India, and from 3.436 in finance and insurance to 3.758 in machinery and equipment. Hence, there are likely to be industry differences. For example, the finance and insurance, utilities, and telecommunications industries used to be heavily regulated and are still frequently perceived as poorly governed. This fact holds particularly true for banking in the aftermath of the financial crisis, even though the respondents themselves are from that same industry. Similarly, the means for innovation and expertise are lower for these industries, which may have similar causes.

The differences in means for the marketing-mix instrument perceptions across countries and industries are somewhat more pronounced than are those for brand strength. As the data shows different profiles across the subsamples, the means themselves are all within their respective standard deviations.

Hence, the sample seems to indicate similar, potentially generalizable patterns of drivers of brand strength across all subsamples. However, the variation around means is substantial and thus we must determine to what extent drivers may yield a general influence and to what extent effects specific to country, industry, or firm moderate them.

4 Estimation and Findings

4.1 Model and Estimation

We apply a linear model in line with Torres et al. (2012). We define brand strength in B2B markets of respondent k's provider j as follows:

$$BS_{jk} = \beta_0 + \beta_1 SC_{jk} + \beta_2 IE_{jk} + \gamma_1 PD_{jk} + \gamma_2 D_{jk} + \gamma_3 PR_{jk} + \gamma_4 C_{jk} + \delta_1 ROLE_k +$$

$$\delta_2 BR_k + \delta_3 SR_k + \sum_{s=1}^{6} \delta_{4+s} D_{IND_{ks}} + \sum_{l=1}^{2} \delta_{10+l} D_{COU_{kl}} + \varepsilon_{jk} \qquad (12)$$

with error term:

$$\varepsilon_{jk} \sim N(0; \sigma^2)$$

BSjk = brand strength of provider j of respondent k's firm

SCjk = sustainability and corporate governance of provider j of respondent k's firm

IEjk = innovation and expertise of provider j of respondent k's firm

PDjk = product performance of provider j of respondent k's firm

Djk = distribution performance of provider j of respondent k's firm

PRjk = price performance of provider j of respondent k's firm;

Cjk = communication performance of provider j of respondent k's firm

ROLEk = role of respondent k (zero equals taking action within the decision process within a group / categories one and three of table A1, one equals taking action within the decision process alone / categories two and four of table A1)

BRk = business reach of respondent k's firm (one equals globally, two equals mostly within home continent, three equals mostly within home country, four equals within specific region of home country)

SRk = share of revenue from B2C compared to B2B activities of respondent k's firm (one equals almost exclusively B2B, two equals majority from B2B, three equals equal share B2B to B2C, four equals majority B2C, five equals almost exclusively B2C)

FSk = Firm size by the number of employees of respondent k's firm (one equals 100 to 250 employees, four equals 250-999 employees, 20 equals 1,000 to 4,999 employees, 40 equals 5,000 to 9,999 employees, 60 equals more than 10,000 employees)

DINDks   = dummy variable for the industry s for respondent k's firm (s equals one: chemicals, commodities, and basic materials; s equals two: information technology; s equals three: telecommunications; s equals four: logistics; s equals five: finance and insurance; s equals six: utilities; reference group: machinery and equipment)

DCOUkl   = dummy variable for the country l for respondent k's firm (l equals 1: Germany; l equals 2: India; reference group USA);

j, k, s, l   = indexes for observation j, respondent k, industry s and country l

The β-, γ-, and δ-coefficients are the parameters to be estimated. The labeling of the parameters reflects our framework. In particular, the β-coefficients denote the coefficients for the brand associations, γ-coefficients those for the marketing-mix perceptions, and δ-coefficients those for the control variables.

As specified above, the two brand association dimensions, the four marketing-mix perceptions, and our control variables explain brand strength. Brand strength is represented by its average scale value. We include brand associations, i.e., 'sustainability and corporate governance' and 'innovation and expertise', as standardized factor values. For marketing-mix performance perceptions, price, product, distribution, and communication, we use averages across items. Controls for the role of respondent within his/her firm are included as a dummy variable, the business reach of the respondent's firm, and the share of revenue from B2C compared to B2B activities as specified above. Additional dummy variables measure fixed effects for industries and countries.

We estimate a sequence of 11 models by applying OLS. Our baseline model I pools all observations across countries and industries. In addition, we estimate separate country-specific models for Germany (II), India (III), and the US (IV). They are built like our baseline

model I, but without country-specific fixed effects. We also estimate industry-specific models for chemicals, commodities, and basic materials (V), information technology (IT) (VI), telecommunications (VII), logistics (VIII), machinery and equipment (IX), finance and insurance (X), and utilities (XI). These are structured like model I without industry-specific fixed effects.

4.2 Fit and Robustness

We find no evidence for harmful collinearity. This finding is supported by the low variance inflation factors for all variables, which are well below the common threshold value of ten (Belsley 1991). Nevertheless, correlations are high for the marketing-mix instrument perceptions (all details are provided in table B1 in appendix B); performance perceptions for the marketing mix jointly reflect a firm's strategic marketing abilities. Hence, we test additional models by excluding variables of the marketing mix while holding the other independent variables constant. The coefficient estimates of the brand associations and the order of magnitude of the coefficients of the marketing-mix perceptions do not change substantially. Thus, we judge our results to be robust against multicollinearity.

In addition, we test alternative model specifications. For all models, we apply both a multiplicative and a semi-logarithmic structure. Our results show only marginal differences in the coefficients, but a slightly inferior goodness-of-fit. We also test nonlinear specifications for the communication perception. Again, while the differences remain marginal, the fit becomes slightly worse. Hence, the model itself seems to be robust, and thus we choose it for the final model.

Since survey data is subject to certain biases, we also test our models with a subsample that includes only those observations for which the respondents spent between ten and 45 minutes for the questionnaire (see Homburg, Klarmann, and Schmitt (2010) for a similar approach). Again, the results stay robust with minor differences.

Finally, we check for a potential common-method bias. We apply the Harman's Single Factor Test (Harman 1976) to our data but do not find evidence for a common-method bias.

4.3 Empirical Results

Table 6 reports our main results for the influences of brand associations and marketing-mix instrument perceptions on brand strength perception for models I through XI.

4.3.1 Main Model

We first analyze the empirical findings of our main model. The explanatory power of our framework is high with an adjusted R^2 of 0.605. We also note that our focal independent variables indicate significant effects, while all control variables except the role of the respondent ($\delta_{1,I}=-0.055$, $p<0.01$) are insignificant. Hence, we can conclude that our framework explains differences in brand strength quite well. We do not find significant differences across firm sizes or industries, as reflected by firm size- and industry-specific fixed effects. However, while India shows no difference from the US in our main model, brand strength in Germany is significantly higher than in the U.S according to its country-specific fixed effects ($\delta_{11,I}=0.105$, $p<0.01$).

Both brand association dimensions show significant positive effects on brand strength. In particular, 'sustainability and corporate governance' ($\beta_{1,I}=0.112$, $p<0.01$) and 'innovation and expertise' ($\beta_{2,I}=0.103$, $p<0.01$) exhibit similar effects, thus emphasizing that both dimensions are important drivers of brand strength in B2B markets. Firms that intend to strengthen their brands should therefore ensure that long-term investments in structures, processes, and systems incorporate aspects of CSR and sustainability. A firm-wide innovation orientation and the nurturing of expertise and creativity appear to be comparably promising avenues for developing strong corporate brands.

Table 6: Regression Results (Dependent Variable: Brand Strength)

Item		Model I (Overall)	Model II (Germany)	Model III (India)	Model IV (USA)	Model V (Chemicals, commodities, and basic materials)	Model VI (IT)	Model VII (Telecommunications)	Model VIII (Logistics)	Model IX (Machinery and equipment)	Model X (Finance and Insurance)	Model XI (Utilities)
Intercept (β_0)		1.236*** (0.109)	1.711*** (0.199)	1.640*** (0.184)	0.699*** (0.175)	0.790*** (0.211)	1.862*** (0.267)	1.324*** (0.273)	1.249*** (0.273)	.875*** (0.265)	.729** (0.356)	1.779*** (0.415)
Sustainability and Corporate Governance (β_1)		0.112*** (0.016)	0.110*** (0.028)	0.167*** (0.026)	0.091** (0.028)	0.089*** (0.033)	0.139*** (0.041)	0.179*** (0.038)	0.120*** (0.042)	-0.022 (0.040)	0.094* (0.053)	0.164*** (0.058)
Innovation and Expertise (β_2)		0.103*** (0.014)	0.135*** (0.027)	0.122*** (0.024)	0.067** (0.025)	0.075** (0.033)	0.172*** (0.035)	0.097*** (0.034)	0.078** (0.038)	0.104*** (0.038)	0.021 (0.047)	0.199*** (0.056)
Product (γ_1)		0.293*** (0.028)	0.247*** (0.052)	0.309*** (0.042)	0.272*** (0.052)	0.261*** (0.064)	0.251*** (0.070)	0.424*** (0.071)	0.173** (0.067)	0.394*** (0.066)	0.360*** (0.090)	0.281*** (0.100)
Distribution (γ_2)		0.254*** (0.026)	0.266*** (0.046)	0.139*** (0.040)	0.388*** (0.049)	0.298*** (0.059)	0.257*** (0.066)	0.079 (0.062)	0.358*** (0.063)	0.337*** (0.067)	-0.037 (0.087)	0.392*** (0.095)
Price (γ_3)		0.100*** (0.023)	0.088** (0.043)	0.083** (0.035)	0.101** (0.041)	0.137** (0.059)	0.074 (0.057)	0.133** (0.055)	0.114** (0.054)	-0.019 (0.056)	0.133** (0.061)	0.111 (0.095)
Communication (γ_4)		0.058** (0.024)	0.015 (0.044)	0.060 (0.037)	0.097** (0.048)	0.099* (0.056)	-0.060 (0.056)	0.038 (0.059)	0.067 (0.067)	0.113* (0.061)	0.334*** (0.077)	-0.194* (0.098)
Role of Respondent (δ_1)		-0.055** (0.025)	0.029 (0.046)	-0.103*** (0.038)	-0.124*** (0.046)	0.013 (0.053)	-0.042 (0.061)	-0.150** (0.072)	-0.139** (0.062)	-0.154** (0.061)	0.109 (0.076)	0.030 (0.096)
Business Reach (δ_2)		-0.016 (0.014)	-0.037 (0.029)	-0.004 (0.021)	-0.009 (0.023)	-0.006 (0.031)	0.027 (0.037)	-0.016 (0.036)	-0.008 (0.031)	0.010 (0.036)	-0.016 (0.043)	-0.046 (0.057)
Revenue from B2B compared to B2C (δ_3)		0.004 (0.011)	-0.023 (0.019)	-0.004 (0.019)	0.034** (0.017)	0.002 (0.024)	-0.019 (0.025)	0.002 (0.026)	-0.029 (0.026)	0.029 (0.028)	0.013 (0.032)	-0.004 (0.040)
Firm size (δ_4)		-6.272e-006 (0.001)	7.687e-005 (0.001)	4.673e-005 (0.001)	7.275e-005 (0.001)	0.001 (0.001)	0.003* (0.002)	-0.001 (0.002)	-0.001 (0.001)	-0.001 (0.002)	0.003 (0.002)	-0.001 (0.002)
Industry based fixed effects[1]	Chemicals, commodities, and basic materials (δ_5)	-0.028 (0.042)	0.021 (0.084)	0.002 (0.057)	-0.145* (0.084)							
	IT (δ_6)	0.054 (0.042)	0.031 (0.084)	0.107* (0.059)	-0.016 (0.081)							
	Telecommunications (δ_7)	0.033 (0.042)	0.053 (0.086)	0.154*** (0.058)	-0.186** (0.080)							
	Logistics (δ_8)	-0.036 (0.041)	-0.037 (0.084)	0.051 (0.058)	-0.178** (0.077)							
	Finance and Insurance (δ_9)	-0.032 (0.051)	-0.016 (0.083)		-0.166** (0.081)							
	Utilities (δ_{10})	-0.023 (0.051)	-0.063 (0.084)		-0.100 (0.080)							
Country based fixed effects[2]	Germany (δ_{11})	0.105*** (0.031)				0.191** (0.076)	0.029 (0.075)	0.234*** (0.088)	0.157** (0.079)	-0.046 (0.088)	0.078 (0.078)	-0.034 (0.110)
	India (δ_{12})	0.011 (0.033)				0.007 (0.067)	0.023 (0.076)	0.193** (0.093)	0.084 (0.075)	-0.110 (0.083)		
N		1656	540	620	496	266	260	276	278	266	152	158
Adjusted R^2		.605	.539	.600	.680	.671	.607	.626	.611	.611	.632	.533
F-value; d.f.; p-value		141.632; 1637; 0.000	40.434; 523; 0.000	67.424; 605; 0.000	66.647; 479; 0.000	46.007; 253; 0.000	34.363; 247; 0.000	39.345; 263; 0.000	37.230; 265; 0.000	35.722; 253; 0.000	24.569; 140; 0.000	17.318; 146; 0.000

*$p<0.1$; **$p<0.05$, ***$p<0.01$; 1) reference group machinery and equipment, 2) reference group US.

All marketing-mix perceptions display significant effects. Although the somewhat weak effect of communication ($\gamma_{4,I}=0.058$, $p<0.05$) on brand strength seems counterintuitive at first, it is meaningful that performance-related instruments such as product ($\gamma_{1,I}=0.293$, $p<0.01$) or distribution ($\gamma_{2,I}=0.254$, $p<0.01$) are major drivers of brand strength in B2B settings. As noted earlier, business relationships in a B2B context are characterized by a greater degree of professionalism and rationality than are B2C relationships. The strong effect of distribution also emphasizes that personal communication by salespeople is far more important in B2B relationships than is traditional communication. This finding is also in line with Fischer and Albers (2010), who report that selling activities in pharmaceutical markets directed towards physicians are more important than are direct-to-consumer campaigns. Since price perception, as measured in our sample, reflects fairness and good value, the effect of price on brand strength ($\gamma_{3,I}=0.1$, $p<0.01$) emphasizes that B2B customers across countries and industries value suppliers' brands that they associate with good quality at reasonable prices. Our empirical evidence also reflects findings from a time-series analysis by Ataman, van Heerde, and Mela (2010), who report that long-term effects of marketing-mix elements are primarily due to product effects (i.e., 60% of the overall effects), secondarily to distribution (32%), and least of all to advertising (6%) or price discounts (2%).

4.3.2 Country Models

Because of the significant differences of brand strength levels between Germany and the US, we look at separate models for the three countries represented in our sample. Using separate models enables us to detect potential differences between effects of our focal variables across nations. As in our main model, the goodness-of-fit of all three country-specific models is very high, with adjusted R^2 values of 0.539 (Germany), 0.6 (India), and 0.68 (US). We note that the base levels of brand strength deviate strongly across countries. We observe base-level brand strength values for Germany and India far above average, but our base level for the

dependent variable is significantly lower for the US, as reflected by intercepts of 1.711 for Germany, 1.64 for India, and 0.699 for the US (all intercepts are highly significant at $p<0.01$). Both effects of the two dimensions of brand associations are highly significant across countries. Nevertheless, we find the most pronounced effect for 'sustainability and corporate governance' for the subsample of India ($\beta_{1,III}=0.167$, $p<0.01$), and that 'innovation and expertise' has the strongest effect for Germany ($\beta_{2,II}=0.135$, $p<0.01$). Although the latter effect on brand strength might reflect that the overall brand strength perception in Germany is driven by technology leadership and innovation, the strong association of sustainability and corporate governance in India could indicate that customers in an emerging market value a long-term orientation.

Similar to our main model, price and product perceptions show comparable and significant effects for all three subsamples. For the US subsample, distribution perception exhibits a far above average effect on brand strength ($\gamma_{2,IV}=0.388$, $p<0.01$), while this impact is substantially lower for India ($\gamma_{2,III}=0.139$, $p<0.01$). Most interesting is the fact that for the subsamples of Germany and India, communication perception reveals no significant association, but we observe a significant effect of communication on B2B brand strength for the US ($\gamma_{4,IV}=0.097$, $p<0.05$).

Although they are not significant in our main model, we detect four significant industry-specific fixed effects for differences of brand strength levels between industries. In India, where telecommunications is a relatively young industry with a few major players, brand strength is significantly higher for this industry than for the reference industry, i.e., machinery and equipment ($\delta_{7,III}=0.154$, $p<0.01$). But in the US, where the telecommunications industry is mature, the brand strength of telecommunications firms is lower than in the reference group ($\delta_{7,IV}=-0.186$, $p<0.05$). We note that the perceived level of brand strength for finance and insurance is also significantly lower ($\delta_{9,IV}=-0.166$, $p<0.05$). Our interpretation is that this signif-

icant difference is a consequence of the financial crisis in 2009 in the US, which is still affecting our sample collected in August and September 2012.

4.3.3 Industry Models

Since we find four significant and two weakly significant industry-specific fixed effects in our country-specific models, we look at industry-specific models to check for different effects of our focal explanatory variables between industries. The overall explanatory power of these models across the seven industries is similar to the main model and the three country-specific models, with adjusted R^2's between 0.533 (utilities/model XI) and 0.671 (chemicals, commodities, and basic materials/model V), which again reflects an excellent fit. We find common support for the positive effects of brand association and marketing-mix instrument perceptions across industries. Therefore, we highlight the most substantial differences.

First, the base levels of brand strength are far above average in the IT and utilities industries, as reflected in pronounced intercepts of 1.862 and 1.779. We find much lower base levels for chemicals, commodities, and basic materials (0.790), machinery and equipment (0.875), and finance and insurance (0.729). Our primary explanation for these base-level differences is that firms in the IT and utilities industries invested heavily in corporate brands in the past; among other reasons, they chose to differentiate themselves after deregulation or as a consequence of shakeouts and market concentration. But the industries with somewhat low levels of brand strength either focus on strong customer relationships and the supplier's reliability (chemicals; machinery) or struggle with consequences of a general crisis of the industry (finance).

Second, 'innovation and expertise' does not show a significant effect on brand strength in the finance and insurance industry. This lack of effect can be explained by low levels of innovation in banking and the difficulty of sustaining the competitive advantages of financial innovations. We also observe that 'sustainability and corporate governance' is not significant in

the machinery and equipment industry. Our interpretation is that in this industry, sustainability is considered as a hygiene factor. This function is also reflected in far above average scores of this construct (see table 5).

Third, we find an additional interesting pattern of effects between our marketing-mix perceptions of distribution and communication. In the IT, the finance and insurance, and the utilities industries, both elements reveal contrary effects. Communication perception shows no effect for the IT and a slightly counterproductive effect for utility industry. These two industries demonstrate far above average base levels of brand strength. Our interpretation is that there is a kind of supersaturation or overspending in these industries. Particularly in power utilities, many firms tried to differentiate their products and services by investing heavily in communication. In the finance and insurance sector, communication displays an highly significant effect on brand strength ($\gamma_{4,X}$=0.334; p<0.01). It is the only highly significant effect of all 11 models. One explanation is that if brands are not thought to be very strong, then communication can still generate higher levels of brand strength. This fact is indicated by the low intercept. A similar effect also holds for the chemicals industry. A second interpretation is that both distribution and 'innovation and expertise' cannot be employed to differentiate oneself from competition, so communication might also serve as a mean of last resort to increase brand strength.

5 Discussion of Empirical Results

In this study, we analyze potential drivers of brand strength in a B2B context. Research on the development of strong brands in B2B markets is critical as B2B brands gain importance, but there are very few empirical studies that address the underlying generating processes and drivers. By using a unique data set comprising three countries and seven industries, we identify and generalize findings across key drivers of B2B brand strength. Thus, we contribute to both academic and managerial knowledge in several ways.

5.1 Implications for Research

Our study provides several implications for researchers. First, it demonstrates that the concept of brand strength established for B2C markets is transferable to B2B markets. Organizational buyers perceive products and services in terms of tangible and intangible attributes. A B2B customer evaluates a B2B brand based on specific associations of intangible attributes. Second, brand strength also reflects customer-based brand equity. This result indicates that the concept of customer-based brand equity is properly applicable to B2B brands.

Third, our study contributes to the understanding of brand strength in an international context. A major difference between B2C and B2B brands is that B2B brands often face a larger group of potential stakeholders, e.g., direct customers, end customers, regulators, employees, suppliers, society, etc. Hence, conceptual differences between the key drivers of brand strength in B2B and B2C markets are likely. Our results provide support for such crucial distinctions. They show that 'sustainability and corporate governance' and 'innovation and expertise' are relevant perceptional drivers of B2B brand strength. As indicated above, these dimensions are quite different from important dimensions in B2C markets, such as excitement, sincerity, and sophistication. This result is essential and generalizable, because it holds across three different countries and seven industries. This finding suggests that brand strength and customer-based brand equity, albeit not fundamentally different from consumer markets, need an augmented conceptual foundation that helps to explain which dimensions are relevant for building brands in B2B markets, and why. A productive avenue for developing a theoretically well-founded conceptualization of B2B brand strength is to start with the relevance dimensions of B2B brands. Fischer, Völckner, and Sattler (2010) suggest the concept of brand relevance in category and brand functions for the B2C markets. According to their findings, which they derive from a cross-national study, risk reduction and social demonstrance are the key benefits brands generally provide to consumers across various product

categories. However, more research is needed to identify whether and to what degree these dimensions and concepts are useful for explaining brand strength in B2B markets.

Fourth, all four marketing-mix perceptions exhibit profound and differential effects on brand strength. While product and distribution perceptions have a major impact, communication perceptions are mostly weak. This result might reflect the fact that in B2B markets, traditional communication instruments are less effective than they are in B2C markets. Target groups are highly heterogeneous, small in size, and often interested only in a limited, specific area of a firm's competence, whereas traditional B2C communication strategies primarily address broad audiences. We speculate that the high impact of product and distribution performance on brand strength is a consequence of complex sales processes that require a high degree of information exchange. Often, sales representatives and front-office personnel provide necessary information to the customers. Thus, both groups might be a promising avenue in B2B markets for nurturing brand strength. We believe it is important to understand that the operational measures reflected in our items are likely to be very different for building a B2B brand compared to a B2C brand. Our study provides support for this hypothesis.

5.2 Implications for Management

Brand associations as well as the product and distribution perceptions prove to be very relevant for the success of a B2B brand. Hence, B2B customers not only value desirable product performance, but also reward emotional attributes that differ from those of B2C brands. For managers, these subjective criteria imply that to focus exclusively on optimizing a firm's purchase funnel by trying to meet product and distribution needs is not sufficient for a brand's success. Instead, those firms will succeed that understand how to serve customers' key requirements while simultaneously linking the relevant associations to their brands.

One crucial touchpoint is the contact between employees and customers. Distribution and product performance are those marketing-mix instruments that matter greatly. I.e., a firm's

brand success depends significantly on the interactions between firms and customers throughout the selling process. Brand strength is effectively induced by investing in professional front-office staff, as reflected in our items within the distribution scale. Investments in this case do not only mean spending financial and personal resources, but also demand that managers recognize employees as a key stakeholder group. Only if a supplier's employees internalize the desired values and clearly communicate them to customers will a firm be able to link brands to desired associations.

We also show that, unlike its importance in B2C markets, the firm's communication strategy is less important in B2B markets. Nevertheless, intangible associations loom large for customers. Hence, communication strategies must diverge from those used in B2C markets. To be effective, firms might de-emphasize major communication campaigns and focus instead on various diversified smaller programs that address different touchpoints between a supplier's B2B brands and its stakeholders.

As argued above, there is a substantial degree of variation among individual stakeholders, e.g., even between stakeholders within different countries and industries. However, key brand associations vary little in their significance, but differ greatly in their effect size across countries. Effects might be even more complex if B2B customers integrate other stakeholders' brand evaluations into their own considerations, as, e.g., in component branding processes. We argue that only by using a consistent positioning in addressing multiple stakeholders will a firm generate favorable brand associations. However, to develop such positioning, firms must prioritize stakeholder groups. Otherwise, firms run the risk that their brands will be perceived as contradictory, which might generate fuzzy associations.

6 Summary

In this contribution, we underline the importance of B2B branding as a field of growing interest for both practitioners and researchers. Although there is a need for more research in

all areas of B2B marketing, we identify brand management as one of the most promising fields for academic research. While it is widely accepted that brands are pivotal in consumer marketing, the role of brand management in B2B marketing has been more or less ignored in marketing research.

Our study helps closing this gap. First, we demonstrate that brand strength is a relevant brand management metric. Second, we recommend that brand associations should be split up into two major drivers of brand strength, 'sustainability and corporate governance' and 'innovation and expertise'. Third, we find strong variation in the effect sizes for marketing-mix instrument perceptions on brand strength: Distribution and product performance are the dominant instruments whereas communication is of subordinate importance. We identify substantial variation in their importance across countries and industries. We find a generalizing link between brand strength and brand associations across countries and industries. Nevertheless, again there is considerable variation in the effects sizes between countries and industries.

Our study has some limitations, and thus suggests some avenues for further research. First, we do not test to what degree B2B brands can have different effects on each of a firm's stakeholder groups. Differences in these effects need further investigation. Second, though we cover the conditions and drivers of successful B2B brands, we still have no evidence on how they might differ from B2C brands in particular. Thus, it is up to future research to compare both in a more detailed way. Third, we use survey data, which is by definition prone to subjective biases of the respondents. Our analysis shares a limitation common to all survey-based models: since it uses cross-sectional data, casual inferences are neither justified nor warranted. We caution that our results should be interpreted while keeping this limitation in mind. As a consequence, we encourage studies that are not based on surveys on B2B branding, since such research might contribute to improved validity, reliability, and objectivity.

REFERENCES PROJECT 2

Aaker, David A., and Robert Jacobson (2001). "The Value Relevance of Brand Attitude in High-Technology Markets," *Journal of Marketing Research*, 38 (4), 485–493.

American Marketing Association (2013). *AMA Dictionary* [available at http://www.marketingpower.com/_layouts/dictionary.aspx; accessed at August 12, 2013].

Ataman, M. Berk, Harald J. van Heerde and Carl F. Mela (2010). "The Long-Term Effect of Marketing Strategy on Brand Sales," *Journal of Marketing Research*, 47 (5), 866-882.

Bureau of Economic Analysis (2013). *Gross-Domestic-Product-(GdP)-by-Industry-Data* [available at http://www.bea.gov/industry/gdpbyind_data.htm; accessed at August 12, 2013].

Belsley, David A. (1991). *Conditioning Diagnostics: Collinearity and Weak Data Regression*, New York: Wiley.

Brown, Tom J., and Peter A. Dacin (1997). "The Company and the Product: Corporate Associations and Consumer Product Responses," *Journal of Marketing*, 61 (1), 68–84.

Churchill, Gilbert A. (1979). "A Paradigm for Developing Better Measures of Marketing Constructs," *Journal of Marketing Research*, 16 (1) 64–73.

Erdem, Tülin, Joffre Swait, and Jordan Louviere (2002). "The Impact of Brand Credibility on Consumer Price Sensitivity," *International Journal of Research in Marketing*, 19 (1), 1–19.

Fischer, Marc (2007). *Valuing Brand Assets: A Cost-Effective and Easy-to-Implement Measurement Approach*, Cambridge: Marketing Science Institute.

Fischer, Marc and Alexander Himme (2013). "20 Years of Brand Valuation: Facts and Insights," Presentation at the *MASB Winter Board Meeting and Summit*, February 14-15, Las Vegas.

Fischer, Marc and Sönke Albers (2010). "Patient- or Physician-Oriented Marketing: What Drives Primary Demand for Prescription Drugs?," *Journal of Marketing Research*, 47 (1), 193-121.

Fischer, Marc, Franziska Völckner, and Henrik Sattler (2010). "How Important Are Brands? A Cross-Category, Cross-Country Study," *Journal of Marketing Research*, 47 (5), 823–839.

Fischer, Marc, Wolfgang Giehl, and Tjark Freundt (2011). "Managing Global Brand Investment at DHL," *Interfaces*, 41 (1), 35–50.

Ghosh, M. and G. John (2009). "When Should Original Equipment Manufacturers Use Branded Component Contracts with Suppliers," *Journal of Marketing Research*, 46 (5), 597–611.

Harman, Harry H. (1976). *Modern Factor Analysis*, Chicago et al.: University of Chicago Press.

Hoeffler, Steve, and Kevin L. Keller (2003). "The Marketing Advantages of Strong Brands," *Journal of Brand Management*, 10 (6), 421–445.

Homburg, Christian, Martin Klarmann, and Jens Schmitt (2010). "Brand Awareness in Business Markets: When is It Related to Firm Performance?," *International Journal of Research in Marketing*, 27 (3), 201–212.

Homburg, Christian, Martin Klarmann, Martin Reimann, and Oliver Schilke (2012). "What Drives Key Informant Accuracy?," *Journal of Marketing Research*, 49 (4), 594-608.

Interbrand (2013). *Best Global Brands 2012*, [available at www.interbrand.com; accessed at April 1, 2013].

Kalra, Ajay and Ronald C. Goodstein (1998). "The Impact of Advertising Strategies on Consumer Price Sensitivity," *Journal of Marketing Research*, 35 (2), 210-224.

Kapferer, Jean N. (2008). *The New Strategic Brand Management*, London et al.: Kogan Page.

Keller, Kevin L., and Donald R. Lehmann (2006). "Brands and Branding: Research Findings and Future Priorities," *Marketing Science*, 25 (6), 740–759.

Lehmann, Donald R., and Yigang Pan (1994). "Context Effects, New Brand Entry, and Consideration Sets," *Journal of Marketing Research*, 31 (3), 364–374.

Monga, Alokparna B., and Loraine Lau-Gesk (2007). "Blending Cobrand Personalities: An Examination of the Complex Self," *Journal of Marketing Research*, 44 (3), 389–400.

Peterson, Robert A. (1994). "A Meta-Analysis of Cronbach's Coefficient Alpha," *Journal of Consumer Research*, 21 (2), 381–391.

Rühle, Alexander, Franziska Völckner, Henrik Sattler, and Claudia Hatje (2012). "Attitude-Based Versus Choice-Behavior-Based Success of Brand Extensions," *Schmalenbach Business Review*, 64 (2), 125–140.

Sethuraman, Raj, Gerard J. Tellis, and Richard A. Briesch (2011). "How Well Does Advertising Work? Generalizations from Meta-Analysis of Brand Advertising Elasticities," *Journal of Marketing Research*, 48 (3), 457–471.

Stahl, Florian, Mark Heitmann, Donald R. Lehmann, and Scott A. Neslin (2012). "The Impact of Brand Equity on Customer Acquisition, Retention and Profit Margin," *Journal of Marketing*, 76 (4), 44–63.

Torres, Anna, Tammo H. A. Bijmolt, Josep A. Tribó, and Peter Verhoef (2012). "Generating Global Brand Equity through Corporate Social Responsibility to Key Stakeholders," *International Journal of Research in Marketing*, 29 (1), 13–24.

Wiersema, Fred (2012). *The B2B Agenda: The Current State of B2B Marketing and a Look Ahead*, University Park: Institute for the Study of Business Markets.

World Bank (2013). *Gross Domestic Product Ranking Table* [available at: http://data.worldbank.org/data-catalog/GDP-ranking-table; accessed August 12, 2013].

Zablah, Alex R., Brian P. Brown, and Naveen Donthu (2010). "The Relative Importance of Brands in Modified Rebuy Purchase Situations," *International Journal of Research in Marketing*, 27 (3), 248–260.

Appendix A: Items of the Controls

Table A1: Items of the Controls

Construct	Item	Response Categories
Role of respondent	For each category your company purchases products/services in, which of the following statements best describes your role in the decision making process?	1 = final decision maker, 2 = part of a decider team, 3 = alone making recommendations, 4 = part of a recommendation team
Business reach of respondents' firm	How would you describe the reach of your business?	1 = globally, 2 = mostly within home continent, 3 = mostly within home country, 4 = within specific region of home country
Share of revenue from B2C compared to B2B activities	How much of your company's revenue derives from B2B compared to B2C activities?	1 = almost exclusively B2B, 2 = majority from B2B, 3 = equal share B2B to B2C, 4 = majority B2C, 5 = almost exclusively B2C

Appendix B: Correlations between the Constructs

Table B1: Correlations between the Constructs

Construct	Brand Strength	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.
Brand Strength	1										
1. Sustainability and corporate governance	0.474***	1									
2. Innovation and expertise	0.415***	0.000	1								
3. Product	0.719***	0.487***	0.453***	1							
4. Distribution	0.695***	0.500***	0.372***	0.738***	1						
5. Price	0.587***	0.538***	0.247***	0.641***	0.658***	1					
6. Communication	0.618***	0.491***	0.374***	0.722***	0.692***	0.614***	1				
7. Role of respondent	0.001	0.095***	-0.023	0.025	0.024	0.071***	0.058**	1			
8. Business reach	-0.063***	-0.030	-0.092***	-0.063***	-0.042*	-0.045	-0.013	0.009	1		
9. Revenue from B2B vs. B2C	-0.026	0.021	-0.052**	-0.038	-0.019	-0.032	0.003	-0.040	0.217***	1	
10. Firm Size	0.010	-0.046**	0.074***	0.008	-0.019	-0.049	-0.040*	-0.143***	-0.259***	0.019	1
Variance Inflation Factor	-	1.797	1.486	3.113	2.801	2.219	2.547	1.099	1.101	1.170	1.213

*p<0.1, **p<0.05, ***p<.01.

PROJECT 3: THE RELEVANCE OF BRANDS IN B2B MARKETS – AN INTERNATIONAL STUDY ACROSS CATEGORIES

Authors: Eric Michael Lennartz and Marc Fischer

ABSTRACT

In this study, we address how brand relevance arises in B2B markets. Using a large survey with B2B managers across seven product categories in Germany and the US, we show that B2B managers use brands to bundle information. This information efficiency evolves as a second-order construct with two reflective subfunctions, namely risk reduction and imagery benefits. B2B brand relevance itself varies substantially between Germany and the US, but variation across product categories is smaller than on B2C markets. In addition, we show that manager and organizational characteristics moderate the influence of brand functions on B2B brand relevance. However, effects of these characteristics differ between Germany and the US. In particular, the responsibility of a manager within a decision process plays a different role in Germany than in the US.

Key Words: Branding and Brand Management; Business-to-Business Marketing; Survey Method; Brand Relevance; Cross-category Analysis.

1 Introduction

Various studies show that it is a key success factor for companies to create valuable brands in customers' minds (e.g., Stahl et al. 2012). While broad evidence supports this view for Business-to-Consumer (B2C) goods, critics argue that managers in Business-to-Business (B2B) markets decide rationally so that brands have no influence (Brown et al. 2011). However, recent studies show that B2B managers include irrational elements into their decisions (e.g., Davis-Sramek et al. 2009). In fact, they find that B2B brands may strongly increase profits (e.g., Fischer, Giehl, and Freundt 2011).

In this context, brand relevance is a key factor for a brand's success. Only if brands are relevant, favorable brand associations can transfer into desired outcomes (Keller 1993). Hereby, business cultures and brand associations largely differ across product categories and countries (e.g., Lennartz et al. 2015; Ronen and Shenkar 2013; Smith, Peterson, and Schwartz 2002). Thus, companies need to track B2B brand relevance across countries and product categories in order to effectively allocate brand investments.

Importantly, as processes in B2B markets fundamentally differ from those processes in B2C markets (e.g., Lemke, Clark, and Wilson 2011; Pick and Eisend 2014), we cannot simply use a B2C-based measurement. We need a B2B-specific tool that allows a valid and easy measurement of brand relevance across product categories and countries. Therefore, our first research aim is to come up with such a measurement. We develop a psychological scale to measure B2B brand relevance. We validate this scale across seven product categories in Germany and the US. Thus, we are the first to measure B2B brand relevance across product categories and countries.

In addition, extensive knowledge exists concerning brand relevance in B2C markets (e.g., Fischer, Voelckner, and Sattler 2010). However, B2B purchase decisions differ strongly from those in B2C markets. They are institutionalized, often involve far more information, and

have organizational and personal components. Particularly, decision criteria and associations with strong brands vary (Lemke, Clark, and Wilson 2011; Wuys, Verhoef, and Prins 2009; Lennartz et al. 2015). Likewise, factors driving the relevance of brands should differ between B2C and B2B markets and across countries and product categories. Nonetheless, they are pivotal to understand why customers use a brand. They bear the underlying factors for successful brand alignments and strategies. Thus, in a second step, we develop and validate a B2B based concept of brand relevance across these categories and countries. We assess antecedents of brand relevance, their proneness to manager and organizational characteristics, and how the impact of these antecedents varies between Germany and the US.

Our study contributes to management and research. From a methodological perspective, we develop a scale for B2B brand relevance. Contrary to existing B2C measures, it covers organizational influences. Researchers can apply this scale to shed light on B2B branding processes. Managers can use it to get an intuition on the profitability of B2B branding strategies across categories and countries. High brand relevance in a category and country may promise revenues from brand investments in the short-term, while low brand relevance may ask for long lasting investments to develop brand relevance.

From a substantive point, we unveil how brand relevance evolves in B2B markets across product categories. We develop and validate a B2B-specific conceptual framework of brand relevance. We show that brand functions and their structure differ from B2C markets. In contrast to B2C markets, brands have only one major function which is to bundle information. This information efficiency evolves as a second-order construct with two reflective subfunctions, namely risk reduction and imagery benefits. Our concept holds for Germany and the US on a general level. But, when we assess the influence of manager and organizational characteristics, we find interesting differences between both countries. Especially, the influence of

the responsibility of a manager within a decision process plays a different role in Germany than in the US.

We structure the remainder of the article as follows: First, we provide a literature review on brand relevance and its meaning for B2B markets. Afterwards, we develop our conceptual framework and describe our questionnaire and sample. Subsequently, we validate our conceptualization, assess variation in brand relevance and functions, analyze antecedents of brand relevance, and discuss them. Finally, we derive managerial implications and draw a conclusion.

2 Review of Existent Studies

Table 1 gives an overview of relevant studies in the literature and compares them to our study. Fischer, Voelckner, and Sattler (2010) define brand relevance as a weight that consumers allocate to a brand in product purchase processes. Although this definition also applies to B2B markets, they only analyze B2C markets. Here, they develop a scale to measure brand relevance and extensively analyze its drivers. Homburg, Klarmann, and Schmitt (2010) show that brand awareness in B2B markets increases brand performance. Hence, brands seem to actually be relevant in B2B markets.

Along these studies, only two studies focus on B2B brand relevance. Zablah, Brown, and Donthu (2010) link the importance of B2B brands to brand sensitivity (intention to consider brand information for decision purposes), preference, and consciousness. In a further study, Brown et al. (2011) show that perceived purchase risk drives brand sensitivity in B2B markets. Similarly to B2C markets, brand relevance seems to arise from specific brand functions.

114

Table 1: Overview on Related Studies

Criterion	Fischer, Voelckner, and Sattler (2010)	Homburg, Klarmann, and Schmitt (2010)	Zablah, Brown, and Donthu (2010)	Brown et al. (2011)	**This Study**
Measurement of brand relevance					
Approach	psychological scale	-	constant sum	constant sum	psychological scale
Adapted to B2B conditions	✕	✕	✓	✓	✓
Conceptualization					
B2B-specific	✕	✓	✓	✓	✓
Full concept of brand functions	✓	✕	✕	✕	✓
Analysis					
Measurement unit	consumer	individual manager	buying center	buying center	individual manager
Across categories	✓	✕	✕	✕	✓
Across countries	✓	✕	✕	✕	✓
Control for					
Manager characteristics	✕	✕	✕	✕	✓
Organizational characteristics	✕	✓	(✓)	✕	✓

Key to the analysis of B2B brand relevance is its measurement. As table 1 shows, studies so far (e.g., Zablah, Brown, and Donthu 2010) have used constant-sum approaches to measure brand relevance. Although these approaches may foster precision of results, they are fairly complex (Srinivasan and Park 1997), which may be especially challenging in business settings. Due to the organizational perspective, survey evaluations are generally more difficult and managers are often reluctant to participate in surveys (Homburg et al. 2012). Moreover, constant-sum approaches depend on managers' potential other decision criteria. As relevant

criteria may differ between B2B purchase situations (e.g., Wuys, Verhoef, and Prins 2009), they provide valid insights for one product category and country, but lack transferability to other categories and countries. Thus, we develop a Likert scale to measure B2B brand relevance. Such a scale is easy to evaluate and allows a valid and comparable measurement across categories and countries.

Concerning the conceptualization, conditions of B2C markets do not apply for B2B markets. For example, B2B markets involve far more levels of decision, namely the individual, buying center, and company level. Thus, we cannot simply apply the B2C concepts by Fischer, Voelckner, and Sattler (2010) in the B2B context. Furthermore, existing B2B concepts on brand relevance cover the impact of brands' risk reduction potential, but not their imagery benefits (e.g., Brown et al. 2011). However, from a conceptual perspective, intangible brand associations are a key factor in building strong B2B brands (e.g., Lennartz et al. 2015). So we develop a new, B2B-specific concept for B2B brand relevance. We expect that the major function of brands is information efficiency. In contrast to B2C markets, we model this function as a second-order construct with two reflective subfunctions: risk reduction and imagery benefits. Thus contrary to existing studies, we are the first to cover a full concept of brand functions in a B2B-specific context.

In addition, our analysis differs in its measurement unit and scope from current studies on B2B brand relevance. As shown in table 1, current studies that measure B2B brand relevance focus on the buying center as measurement unit. This approach provides useful results. Nevertheless, we expect additional insights to come from the individual manager perspective. For example, risk in organizational contexts does not only refer to the overall organizational risk, but also has a personal component linked to the individual manager (Hawes and Barnhouse 1987). Therefore, in line with Homburg, Klarmann, and Schmitt (2010), our measurement unit is the individual manager. Moreover, as in B2C markets (Fischer, Voelckner, and Sattler

2010), we expect differences across product categories and countries. While existent studies focus on one product category and country, we are the first to include different categories and an additional country to our analysis. Finally, due to their focus on the buying center as measurement unit, existing studies on B2B brand relevance hardly control for individual manager characteristics. At the same time, research provides evidence that brand relevance may vary across manager, company, and country characteristics (e.g., Homburg et al. 2012). Thus, we expect additional insights from including a broader set of manager and organizational characteristics. We address this gap and assess antecedents of brand relevance, while largely controlling for individual manager and organizational characteristics.

3 Conceptual Framework

We follow the framework depicted in Figure 1. B2B brand relevance depends on the need for heuristics that bundle information for a purchase decision (information efficiency). Two subfunctions reflect this process. Managers try to reduce risk that is a consequence of a lack of information (risk reduction) and use brands to get a general impression of imagery components (imagery benefit). In addition, product-market characteristics may drive B2B brand relevance. Manager and organizational characteristics should moderate these processes.

Figure 1: Conceptual Framework for B2B Brand Relevance
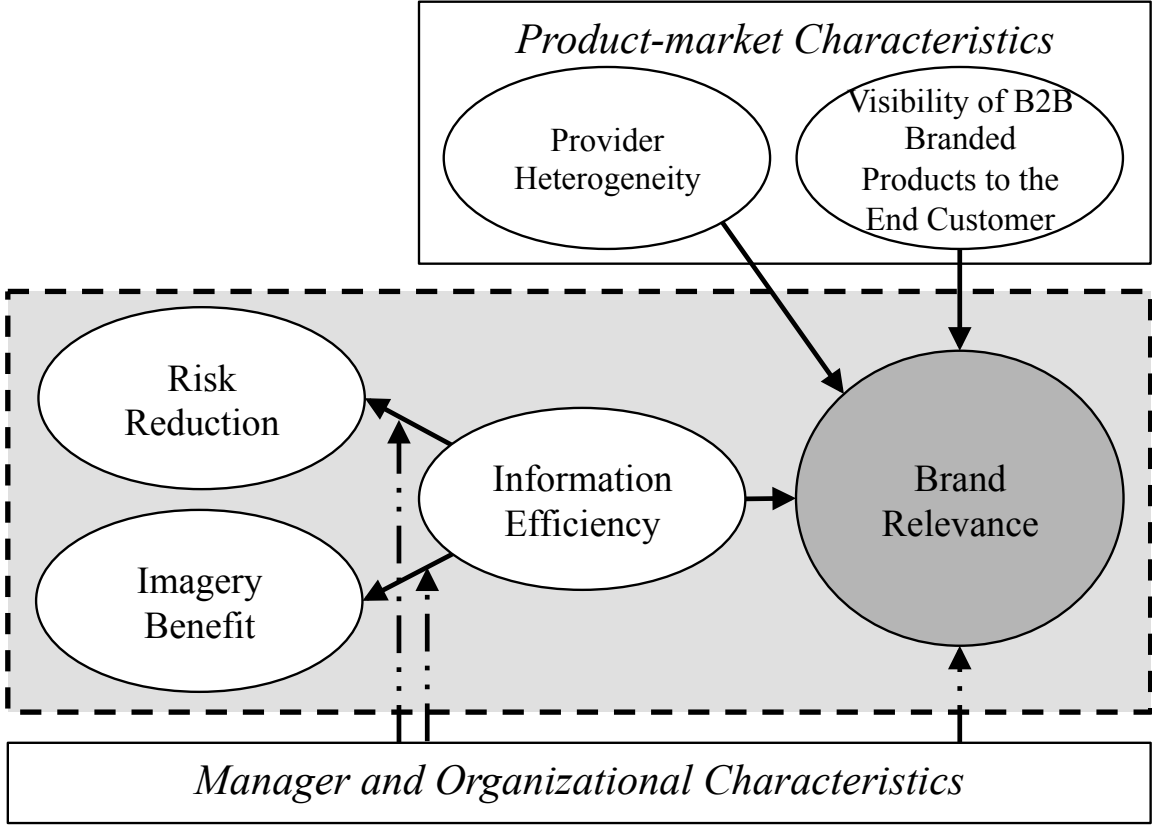
3.1 Information Efficiency as Metafunction

According to information processing theory (e.g., Atkinson and Shiffrin 1968) and resource availability theory (Miller 1956), individuals have limited memory capacities to store, process and retrieve information. Products and services combine various attributes, associations and functions, both immaterial and material. In product evaluation and purchase decisions, consumers have to process information on all of these components. Due to their amount, they simply cannot evaluate all information. They apply heuristics to process information more efficiently (e.g., La, Patterson, and Styles 2009). Brands serve as such heuristics that simplify the choice and its predicted utility (Brown et al. 2011). They bundle information and reflect an overall evaluation.

In B2B markets, buying centers plan and transact purchase processes. While such multi-person structure improves efficiency, buying center members still have limited memory ca-

pacities (Zablah, Brown, and Donthu 2010). Although resources such as information systems help circumvent these limitations, managers finally have to combine all information. Likewise, B2B purchases involve a high complexity and customized conditions, which increases the amount of information and the difficulty to process them for a thorough decision (Brown et al. 2011; Johnston and Lewin 1996). Managers likely lack expertise to rationally process and evaluate information (La, Patterson, and Styles 2009). Here, B2B brands bundle information and ensure *information efficiency*.

3.2 Risk Reduction by B2B Brands

Two subfunctions reflect this overall function. First, each product characteristic may cause a product mistake. Information economics theory (Darby and Karni 1973; Nelson 1970) divides characteristics into search, experience and credence characteristics. While buyers can evaluate search characteristics ex ante to a purchase, they may only evaluate the latter two after a purchase. They thus bear extensive risk. Due to their complexity, B2B products normally have many experience and credence characteristics (e.g., La, Patterson, and Styles 2009).

Moreover, information on many search characteristics connects to high personnel and financial efforts. Complexity and multitude of product components (e.g., Aaker and Jacobson 2001) and a lack in expertise of B2B managers (La, Patterson, and Styles 2009) further lower the utility of information search. If the expected utility to search information outweighs the disutility of search costs, search characteristics become experience or credence characteristics (Darby and Karni 1973). As brands provide intuition on those criteria not assessed ex ante, buying center members may use brands to reduce the risk of a purchase decision.

Contrary to B2C markets, in B2B markets risk splits up into company and personal risk (Hawes and Barnhouse 1987). The company risk is the risk that a wrong purchase decision affects the performance of a company. It may be large when B2B purchases involve high fi-

nancial, temporal and personnel resources (e.g., Luo and Kumar 2013; Wuys, Verhoef, and Prins 2009). Personal risk results from the fact that companies evaluate, reward, and develop managers based on their decisions (Cooper, Wakefield, and Tanner 2006).

3.3 Imagery Benefits of B2B Brands

Moreover, brands are an efficient cue for all information regarding image beliefs (e.g., Keller 1993). They bundle information on associations and interpretations of companies' actions, quality of products sold under this brand, branding contexts etc. (e.g., Brown and Dacin 1997; Cleeren, van Heerde, and Dekimpe 2013). Individuals use them to communicate certain values to themselves and others. Brands enhance individuals' self-concept if they fit the desired self-concept and -beliefs and have a common interpretation across peer groups (Grubb and Grathwohl 1967; Grubb and Hupp 1968).

Due to the institutionalized nature of B2B purchases, researchers often disregard this imagery benefit (e.g., Brown et al. 2011). But final decisions makers are still individuals that communicate their own and a company's self-concept to their environment. Thus, we expect that the imagery benefit function exists in B2B markets, but has a much broader scope than the social demonstration function in B2C markets (Fischer, Voelckner, and Sattler 2010). In addition to a personal self-concept, it may contribute to corporate associations, such as sustainability and corporate governance (Lennartz et al. 2015).

Moreover, B2B brands serve a broader range of stakeholder groups. First, they help enhance the self-concepts of buying center members and managers. They have to represent their self-concepts against various groups including themselves, other buying center members and senior managers. This representation may influence their payment, job satisfaction, or career opportunities.

Second, B2B brands support the purchasing company. Companies' corporate associations and their brands' images evolve from their own actions and related factors (e.g. Brown and

Dacin 1997). Members of companies in the downstream of the supply chain, own employees, and other stakeholders interpret the usage of a brand by a company and develop associations with the company, its products, and services. These associations may also help attract talented employees (DelVecchio et al. 2007). In addition, purchasing brands with a beneficial image may lead to stronger organizational identification because it involves identification with upstream members of the supply chain (Hughes and Ahearne 2010). Subsequently, employees act more customer-oriented and congruent with a company's aims (Löhndorf and Diamantopoulos 2014). Moreover, brands represent employees in their personal communication with business partners.

Third, B2B brands support other B2B or B2C brands in the downstream of the supply chain by fitting to the self-concepts of end customers or the image of companies and their employees (e.g., Yoon, Gürhan-Canli, and Bozok 2006). Similar to the brand of a product, end customers use brands in the upstream of the supply chain to enhance their self-concept. This notion especially holds for ingredient branding processes, where a brand uses brands of components and incorporates consumers' associations to them into its own brand profile (Ghosh and John 2009). Brands that use ingredient branding are for example Intel or Shimano.

3.4 Product Market-specific Antecedents of B2B Brand Relevance

We expect the visibility of a B2B branded product or service to end customers to drive brand relevance. Like for ingredient branding, B2B brands help enhance the brand equity of the "final brand" (Desai and Keller 2002). Their impact increases with their awareness and, thus, their visibility to the end customer (Keller 1993). Moreover, if more heterogeneous providers act in a market, customers may more easily find other criteria than the brand to distinguish products, which leads to a decreasing brand relevance.

These factors are equivalent to the variables visibility of consumption and number of available brands in the conceptualization of Fischer, Voelckner, and Sattler (2010). Contrarily, the factors decision involvement and group decision-making in their conceptualization represent manager or organizational characteristics. We expect them to moderate the effects of brand functions. Hence, we only consider their B2B equivalents as moderators in our conceptual model. All other variables by Fischer, Voelckner, and Sattler (2010) lack relevance in B2B markets, so we do not consider them in our concept.

3.5 Moderation Effects by Manager and Organizational Characteristics

We expect manager and organizational characteristics instead of demographics to moderate relationships in B2B markets. With respect to manager characteristics and in line with literature (Homburg et al. 2012; Lennartz et al. 2015; Zablah, Brown, and Donthu 2010), respondent's responsibility and position in a company hierarchy may be moderators. Regarding organizational characteristics, we identify company size, buying center size, and global reach as potential moderators from related research (Homburg, Klarmann, and Schmitt 2010; Torres et al. 2012). However for all factors, effects may exist in both directions and differ between countries. Hence, we do not formulate hypotheses about the direction of effects ex ante.

Moreover, while the US show a strong service orientation, industry is far more important in Germany. For 2012, the US generated 78% of their GDP in the service sector and 21% in the industry sector, while Germany only generated 69% in the service, but 30% in the industry sector (World Bank 2016). Furthermore, both countries reveal few, but fundamental differences in their culture (Hofstede 2016). Specifically, Germans value long-term orientation more than US-Americans. Concerning business culture, both countries belong to different clusters (e.g., Ronen and Shenkar 2013; Smith, Peterson, and Schwartz 2002). These clusters mainly differ in leadership style which is more humane oriented in the US than in Germany

(Dorfman, Hanges, and Brodbeck 2004). Therefore, we expect effects to vary across countries.

4 Questionnaire and Data Sample

4.1 Focal Constructs

Table 2 shows the scales for our focal constructs brand relevance, risk reduction, and imagery benefit. We measure all items via fully-labeled five-point Likert scales. Respondents answer scales for one out of seven B2B product categories.

Table 2: Items of the Focal Constructs

Scale	Items
B2B brand relevance[1]	1. The strength of a provider's brand influences my purchase decision fundamentally. 2. It is important to me to purchase products or services from a company with a strong brand. 3. I consciously look for companies with a strong brand.
Risk reduction[1]	When purchasing products of (*category*)[2], I choose providers with strong brands 1. … as reassurance of good quality. 2. … as they will remain compatible with future products. 3. … as they will keep offering high-quality extensions. 4. … as they usually have better solutions for our problems. 5. … as they usually provide more individual solutions.
Imagery benefit[1]	When purchasing products of (*category*)[2], I choose providers with strong brands because 1. … they fit well with our company's values. 2. … they demonstrate our corporate success. 3. … our customers value them more. 4. … our customers ask for them. 5. … they positively support the image of our employees.

1) we measure items on a fully labeled Likert scale (1 = strongly disagree, 2 = disagree, 3 = neither agree nor disagree, 4 = agree, 5 = strongly agree); 2) we replace (category) by one of our seven product categories.

For the measurement of brand relevance and brand functions, the B2C scales of Fischer, Voelckner, and Sattler (2010) serve as our starting point. From these scales, we adapt those items that fit to B2B contexts. However, as we expect differences between B2B and B2C constructs, we also develop new B2B-specific items based on our conceptualization and underlying theories. We discuss these items in various focus groups with experts from different industries and apply them to several pre-studies. Based on the results, we delete and adapt items

to ensure reliability and validity. Two five-item scales result. Risk reduction covers the reassurance of quality, compatibility, and customized solutions today and in the future. Imagery benefit covers benefits for a purchasing company, its employees, buying center members, and customers.

4.2 Product-market Antecedents and Moderators

For practicability reasons, we use single item scales to measure product-market antecedents and moderator variables. Single-item scales reduce response biases due to mental fatigue and carelessness (Netemeyer, Bearden, and Sharma 2003). We cover an item for the visibility of B2B branded products to the end customer and the provider heterogeneity. Further, respondents report their responsibility concerning purchase decisions and the number of reporting levels to the CEO. Finally, we ask for their company's industry, size, buying center size and global reach.

4.3 Data Sample

A global market research firm collected the data via an online survey among managers in Germany and the US during August and September 2012. The overall sample size is 721. We exclude 75 respondents with a uniform response style (standard deviation across all items for brand relevance and functions <0.20) and seven respondents who spend less than five minutes on the questionnaire (average response time was 14 minutes).

Table 3 provides an overview of our detailed sample structure. Respondents are decision makers in B2B firms with at least 100 employees in Germany and the US. These countries are highly relevant because they cover about 27.3% of the World's total GDP in 2014 (World Bank 2016). Moreover, we randomly chose firms. Firms come from various industries and cover all industries of the Global Industry Classification Standard (Standard & Poor's 2016). Within the survey, respondents rate items concerning their providers in one out of seven

product categories. We randomly assigned respondents to these product categories. Categories count for more than 30% of the US GDP in 2014 (Bureau of Economic Analysis 2016) and are, thus, highly relevant, too. Overall, this detailed research setting enables us to control for industry, category and country influences.

Table 3: Sample Composition

Respondent characteristics		Distribution across category evaluated by respondent	
Country	*Origin of respondents*	Chemicals, commodities and basic materials	91
Germany	333	Information Technology	98
USA	326	Logistics	94
#obs.	659	Telecommunication	92
Industry	*Origin of respondents*	Machines and components	96
Energy and materials	41	Banking and insurances	93
Aerospace and defense	30	Utilities	95
Automotive and transportation	66	#obs.	659
Information Technology	104		
Telecom and media	28		
Other high tech	34		
Banking and insurance	65		
Other professional services	90		
Pharmaceuticals	22		
Consumer Goods	54		
Other	125		
#obs.	659		

5 Conceptual Validation

As we introduce a new tailored B2B framework of brand relevance, we test its reliability and validity. Table 4 shows detailed reliability measures. All values are above 0.8 for Cronbach's alpha and above 0.5 for the average variance extracted, which confirms the high

internal consistency (e.g., Netemeyer, Bearden, and Sharma 2003). A preliminary exploratory factor analysis displays three factors that represent our intended constructs. In a confirmatory factor analysis, we assess the impact of information efficiency on brand relevance as given in the grey box of figure 1. We model information efficiency as a second-order reflective construct with the subfunctions risk reduction and imagery benefit.

Table 4: Reliability Measures for B2B Brand Relevance, Risk Reduction and Imagery Benefit

	Coefficient alpha	*Average variance extracted*
B2B brand relevance	0.87	0.69
Risk reduction	0.86	0.54
Imagery benefit	0.89	0.63

The confirmatory factor analysis supports our conceptualization. Fit criteria meet required thresholds from literature (e.g., Bagozzi and Yi 1988) and indicate an excellent model fit. All factor loadings are highly significant (all t-values > 17). For the pooled data, indicator reliabilities lie between 0.49 and 0.89, and composite reliabilities range from 0.79 to 0.89. Common fit indices also show that the model suits the data well: the root mean squared error of approximation is 0.05, the standardized root mean square residual is 0.03, the comparative fit index is 0.98, and the Tucker-Lewis-index is 0.98. Our second order construct for information efficiency explains 73.8% of the variance in B2B brand relevance.

For the discriminant validity of risk reduction and imagery benefit, we compare our model to a model where all items reflect one brand function. Separate Chi-square test for our pooled (ΔChi-Square$_{pooled}$=257.90; Δd.f..=1; p<.001), German (ΔChi-Square$_{GER}$=176.15; Δd.f.=1; p<.001), and US sample (ΔChi-Square$_{US}$= 89.50; Δd.f.=1; p<.001) each favor a solution with two subfunctions.

In a two-group structural equation model, we test whether constructs reveal configural, metric, and scalar measurement invariance between Germany and the US. We assume config-

ural measurement invariance because country-wise samples reveal similar patterns of loadings and sufficient values for fit criteria (Bollen 1989). When inducing invariance instructions for partial scalar and metric measurement invariance, information criteria slightly improve (ΔBayesian information criterion=-1.90, ΔAikake information criterion=-0.48) and fit criteria are comparable to those of the original model (RMSEA$_{restricted}$=0.04, SRMR$_{restricted}$=0.03, CFI$_{restricted}$=0.97, TLI$_{restricted}$=0.97). We may, thus, conduct cross-national comparisons (Steenkamp and Baumgartner 1998).

In addition, we check the relation of brand relevance with the willingness to pay a price premium, the repurchase probability, the brand strength perception of the primary provider, and brand importance by a constant sum approach compared to the four marketing-mix dimensions. Table A1 in appendix A illustrates detailed measures. Significant correlations in table 5 underline convergent and nomological validity of our brand relevance scale.

Table 5: Correlation between B2B Brand Relevance and Validity Criteria

Measurement	#items	N	Coefficient alpha	Correlation with B2B brand relevance[1]
B2B brand relevance	3	659	0.87	-
Convergent				
Brand importance	1	659	-	0.29***
Nomological				
Price premium	1	659	-	0.66***
Brand consideration	1	659	-	0.41***
Brand strength of primary provider[2]	3	659	0.82	0.30***

1) correlations have been corrected by reliabilities; 2) represents the average scale value of the three items chosen, see table A1 for an exact definition; ***$p<0.001$.

6 Variation of B2B Brand Relevance and Its Antecedents

6.1 Variation across Countries

Table 6 shows brand relevance and brand functions across categories and countries. All categories reveal latent means greater than 3.50 for the US and 3.20 for Germany. These val-

ues are greater than the midpoint of the initial items and far greater than the scale minimum one which would indicate that respondents disagree on considering brands for purchases. Therefore, brands seem to have substantial relevance in B2B markets.

Table 6: Latent Construct Means for Brand Relevance and Brand Functions by Country and Product Category[1,2]

Countries		Categories					
		Pooled		Germany		USA	
B2B BRiC							
USA	3.66	Machines and Components	3.72	Machines and Components	3.70	Machines and Components	3.75
Germany	3.41 ***	IT	3.58	Logistics	3.44 †	Utilities	3.73
		Logisitcs	3.56	IT	3.43 *	IT	3.72
		Utilities	3.48 *	Banking and Insurances	3.37 *	Logistics	3.68
		Banking and Insurances	3.48 *	Chemicals, Commodities, and Basic Materials	3.33 *	Banking and Insurances	3.59
		Telecommunication	3.44 **	Telecommunication	3.32 *	Telecommunication	3.57
		Chemicals, Commodities, and Basic Materials	3.44 **	Utilities	3.25 **	Chemicals, Commodities, and Basic Materials	3.55
Information Efficiency							
USA	3.63	Machines and Components	3.66	Machines and Components	3.63	Machines and Components	3.70
Germany	3.39 ***	Logistics	3.58	Logistics	3.46	Logistics	3.69
		IT	3.49 †	Telecommunication	3.40	Utilities	3.69
		Banking and Insurances	3.47 †	Banking and Insurances	3.37 †	IT	3.62
		Telecommunication	3.47 *	IT	3.35 *	Chemicals, Commodities, and Basic Materials	3.59
		Chemicals, Commodities, and Basic Materials	3.46 *	Chemicals, Commodities, and Basic Materials	3.32 *	Banking and Insurances	3.58
		Utilities	3.43 *	Utilities	3.19 **	Telecommunication	3.54
Risk Reduction							
USA	3.66	Machines and Components	3.70	Machines and Components	3.68	Machines and Components	3.73
Germany	3.43 ***	Logistics	3.59	Logistics	3.48	Utilities	3.72
		IT	3.53 †	Telecommunication	3.47	Logistics	3.71
		Telecommunication	3.52 †	IT	3.40 †	IT	3.65
		Banking and Insurances	3.49 *	Banking and Insurances	3.39 †	Chemicals, Commodities, and Basic Materials	3.62
		Chemicals, Commodities, and Basic Materials	3.49 *	Chemicals, Commodities, and Basic Materials	3.35 *	Banking and Insurances	3.60
		Utilities	3.46 *	Utilities	3.22 **	Telecommunication	3.57
Imagery Benefit							
USA	3.53	Logistics	3.54	Logistics	3.41	Logistics	3.69
Germany	3.26 ***	Machines and Components	3.48	Machines and Components	3.41	Utilities	3.56
		Banking and Insurances	3.42	Banking and Insurances	3.33	Machines and Components	3.55
		Chemicals, Commodities, and Basic Materials	3.39	Telecommunication	3.29	Chemicals, Commodities, and Basic Materials	3.53
		Telecommunication	3.37	Chemicals, Commodities, and Basic Materials	3.25	Banking and Insurances	3.51
		Utilities	3.29 *	IT	3.12 †	Telecommunication	3.46
		IT	3.28 *	Utilities	3.03 *	IT	3.43 †

1) highest values for country and product category at the top; 2) standard deviations in parentheses; †$p<0.10$, *$p<0.05$, **$p<0.01$, ***$p<0.001$, two-sided t-test to the highest value of a construct in a country and category.

B2B brand relevance is significantly higher in the US ($M_{B2BBrandRelevance}$ = 3.66) than Germany ($M_{B2BBrandRelevance}$ = 3.41). This result goes along with the findings of Fischer, Voelckner, and Sattler (2010) that overall B2C brand relevance is significantly higher in the US than other countries. Similarly, brand functions have higher values in the US than Germany. We explain this result by customer values. In the US, individualism has much higher importance in comparison to Germany (Hofstede 2016). Likewise, brands support managers, but also companies, to differentiate themselves from each other. Individualism may, thus, emphasize brand relevance in B2B purchase decisions in the US compared to Germany. In addition, marketing spending per capita in the US is higher than in Germany (Strategy Analytics 2016). Long-term investments in brands in the US may also have built up brand relevance.

6.2 Variation across Categories

Brands are most relevant for Machines and Components in Germany ($M_{BrandRelevance}$ = 3.70) and the US ($M_{BrandRelevance}$ = 3.75). Overall the ranking is similar between countries. Especially, product categories with rather low complexity, such as Chemicals, Commodities, and Basic Materials and Telecommunication, rank low while complex product categories, such as Machines and Components, IT, and Logistics, rank high. As assumed, potential information overload seems to drive B2B brand relevance.

The category utilities depicts the only deviation in the similarity of country-wise patterns. It ranks lowest in Germany ($M_{BrandRelevance}$ = 3.25) and shows the second highest brand relevance in the US ($M_{BrandRelevance}$ = 3.73). Two notions may explain this difference. On the one hand, Utilities face a much lower degree of liberalization in Germany than in the US. In detail, German companies underlie restrictions in their provider choice. These restrictions emphasize the need to consider legal regulations for purchase decisions. Brand, thus, are less important for these decisions. On the other hand, media reports, green movements, and political decisions have fostered a bad image of utility companies among German customers. As

this image is consistent across companies, brands hardly help differentiate purchase alternatives from each other. They loose relevance for purchase decisions.

The ranking of the imagery benefit function shows some remarkable deviations to the ranking of B2B brand relevance. IT ($M_{Image} = 3.28$) ranks lowest while Logistics ($M_{Image} = 3.54$) ranks highest. This result may reflect the importance of personal contact in B2B markets. While Logistics are highly visible to subsequent customers, IT products may be hardly noticeable.

7 Model for the Antecedents of B2B Brand Relevance

We model B2B brand relevance to depend on risk reduction and imagery benefit instead of using information efficiency. This approach enables us to get a more detailed picture of which brand facets drive brand relevance. We include B2B brand relevance and functions by their latent construct values. Consistent with the B2C model of Fischer, Voelckner, and Sattler (2010), we include product-market specific factors as direct effects. Here, we use the visibility of a B2B branded product or service to the end customer and the provider heterogeneity from our conceptual framework. We operationalize both as respondents' value on a single-item scale as explained above. In detail, we specify a model as follows:

$$
\begin{aligned}
\text{BRANDRELEVANCE}_{ijkm} = \beta_{oi,l} + \beta_{1i,m}\text{RISK}_{ikm} + \beta_{2i,m}\text{IMAGE}_{ikm} \\
+ \gamma_{1,m}\text{VISIBILITY}_{ik} + \gamma_{2,m}\text{HETEROGENEITY}_{ik} \\
+ \vartheta_j + \eta_k + \zeta_{ijkm}
\end{aligned}
\tag{13}
$$

with

$$
\vartheta_j, \eta_k, \zeta_{ijkm} \sim N(0, \Sigma)
$$

where:

BRANDRELEVANCE$_{ikm}$ = B2B brand relevance for respondent i in industry j evaluating category k in country m,

$RISK_{ikm}$	= importance of brands' risk reduction function for respondent i evaluating category k in country m,
$IMAGE_{ikm}$	= importance of brands' imagery benefit function for respondent i evaluating category k in country m,
$VISIBILITY_{ik}$	= visibility of a B2B branded product to end customers perceived by respondent i for category k (1 = not visible, 2 = hardly visible, 3 = visible, 4 = highly visible),
$HETEROGENEITY_{ik}$	= provider heterogeneity in category k as perceived by respondent i (1 = very few, mostly undifferentiated providers to choose from, 2, 3, 4, 5 = large number of clearly differentiated providers to choose from),
β, γ	= (unobserved) parameters,
ϑ_j	= industry-specific error,
η_k	= category-specific error,
ζ_{ijkm}	= idiosyncratic error,
i	= 1, 2, …, I (number of respondents),
j	= 1, 2, …, J (number of respondents' industries),
k	= 1, 2, …, K (number of categories that respondents evaluate),
m	= Germany, USA.

The parameters to be estimated are the β - and γ -coefficients. We also include a category-, industry-specific and idiosyncratic error term. While the category-specific error term covers variance in the error across the categories for which respondents evaluate brand relevance, the industry-specific error term covers variance in the error across the industries that respondents' companies belong to. Error terms are normally distributed with zero mean and variance-covariance matrix Σ .

We model β-coefficients, including the intercept, as random. In this manner, we cover heterogeneity from individual manager and organizational characteristics in the intercept and the effects of the brand functions. Therefore, we split up the β-coefficients in a respondent-independent part ($\overline{\beta_{n,m}}$) and a respondent-biased part and estimate them based on moderators as follows:

$$
\begin{aligned}
\beta_{ni,m} = {} & \overline{\beta_{n,m}} \\
& + \lambda_{n1,m}\text{RESPONSIBILITY}_i + \lambda_{n2,m}\text{REPORTINGLEVELS}_i + \lambda_{n3,m}\text{SIZE}_i \\
& + \lambda_{n4,m}\text{BUYINGCENTERSIZE}_i + \lambda_{n5,m}\text{BUSINESSREACH}_i \\
& + \xi_{nim}
\end{aligned}
\tag{14}
$$

with

$$
\xi_{nim} \sim N(0, \mathbf{T})
$$

where:

RESPONSIBILITY$_i$ = Responsibility of respondent i in B2B purchase decisions (taking within the decision process within a group = 0, taking within the decision process alone = 1),

REPORTINGLEVELS$_i$ = Number of reporting levels of respondent i to the CEO (0 = I am the CEO, 1 = I directly report to the CEO, 2 = 2 levels, 3 = 3 levels, 4 = 4 levels, 5 = 5 or more levels),

SIZE$_i$ = Number of employees of respondent i's company (1 = 100-249, 4 = 250-999, 20 = 1,000-4,999, 40 = 5,000-9,999, 60 = 10,000 and more employees),

BUYINGCENTERSIZE$_i$ = Buying center size in respondent i's company (1 = only me, 2 = me and another person, 3 = me and two other persons, 4 = me and three other persons, 5 = me and more than three other persons),

BUSINESSREACH$_i$	= Global business reach of respondent i's company (1 = globally, 2 = home continent, 3 = home country, 4 = specific region of our home country),
λ	= (unobserved) parameters,
ξ_{nim}	= individual-specific error for coefficient $\beta_{ni,m}$,
i	= 1, 2, …, I (number of respondent),
m	= Germany, USA,
n	= 0 (intercept), 1 (risk reduction function), 2 (imagery benefit function).

The λ-parameters model the impact of manager and organizational characteristics on the β-coefficients. Here, in line with Fischer, Voelckner, and Sattler (2010), we include the moderators of our conceptual framework as interaction terms. Moderators represent those characteristics that may shape the influence of brand functions on brand relevance. We include the responsibility of a respondent as a binary variable that is one when s/he takes actions on her own and zero otherwise. Buying center size and the number of reporting levels to the CEO are count variables. The company size represents the number of employees in a company. Business reach measures how limited a company's target market is to a specific region. We code all moderators as given above. The individual-specific error terms ξ_{nim} account for the individual deviation in the β-coefficients. They are normally distributed with zero mean and variance-covariance matrix **T**.

8 Estimation Results

We insert equation (14) into equation (13) to get our full estimation equation and estimate separate models for Germany and the US via maximum simulated likelihood. We allow free correlations between random parameters. Table 7 shows our estimation results.

Table 7: Regression Results

Dependent: B2B brand relevance	Coefficient[1]	Germany		USA	
Intercept	$\overline{\beta_{0,m}}$	-0.26 **	(0.10)	0.66 **	(0.20)
Standard deviation of individual-specific error		0.77 ***	(0.03)	0.81 ***	(0.05)
x Responsibility	$\lambda_{01,m}$	0.44 ***	(0.05)	-0.47 ***	(0.11)
x ReportingLevels	$\lambda_{02,m}$	-0.07 **	(0.03)	-0.06	(0.05)
x Size	$\lambda_{03,m}$	2.38×10^{-3} †	(1.22×10^{-3})	3.37×10^{-3}	(2.23×10^{-3})
x BuyingCenterSize	$\lambda_{04,m}$	0.01	(0.02)	-0.14 **	(0.04)
x BusinessReach	$\lambda_{05,m}$	0.20 ***	(0.03)	-0.05	(0.05)
Risk Reduction	$\overline{\beta_{1,m}}$	0.64 ***	(0.06)	0.04	(0.11)
Standard deviation of individual-specific error		0.20 ***	(0.01)	0.24 ***	(0.02)
x Responsibility	$\lambda_{11,m}$	-0.24 ***	(0.03)	0.37 ***	(0.06)
x ReportingLevels	$\lambda_{12,m}$	0.03 *	(0.01)	-0.05 †	(0.02)
x Size	$\lambda_{13,m}$	-1.68×10^{-3} *	(0.73×10^{-3})	0.34×10^{-3}	1.42×10^{-3}
x BuyingCenterSize	$\lambda_{14,m}$	0.06 ***	(0.01)	0.19 ***	(0.02)
x BusinessReach	$\lambda_{15,m}$	-0.03 †	(0.02)	0.09 ***	(0.03)
Imagery Benefit	$\overline{\beta_{2,m}}$	0.44 ***	(0.05)	0.75 ***	(0.09)
Standard deviation of individual-specific error		0.05 ***	(0.00)	0.02 ***	(0.00)
x Responsibility	$\lambda_{21,m}$	0.13 ***	(0.03)	-0.25 ***	(0.05)
x ReportingLevels	$\lambda_{22,m}$	-0.01	(0.01)	0.06 **	(0.02)
x Size	$\lambda_{23,m}$	0.74×10^{-3}	(0.62×10^{-3})	-0.36×10^{-3}	(1.35×10^{-3})
x BuyingCenterSize	$\lambda_{24,m}$	-0.07 ***	(0.01)	-0.15 ***	(0.02)
x BusinessReach	$\lambda_{25,m}$	-0.02	(0.02)	-0.08 ***	(0.02)
Visibility	$\gamma_{1,m}$	0.03 ***	(0.01)	0.04 ***	(0.01)
Heterogeneity	$\gamma_{2,m}$	0.02 **	(0.01)	1.74×10^{-3}	(8.25×10^{-3})
Standard deviations of error components					
Product Category-specific error		0.07 ***	(0.00)	0.06 ***	(0.00)
Industry-specific error		0.02 ***	(0.00)	0.09 ***	(0.00)
Idiosyncratic error		0.93×10^{-3} ***	(0.03×10^{-3})	0.81×10^{-3} ***	(0.03×10^{-3})
Number of respondents		280		285	
Log-likelihood		-27.11		9.96	

1) index m indicates the country (Germany, USA); †p<0.10, *p<0.05, **p<0.01, ***p<0.001 based on a two-sided test; standard errors in parentheses.

8.1 Robustness and Fit

The overall fit of our models is good. Pseudo-R^2 is 0.90 each for Germany and the US.

Results stay similar when we estimate a model with information efficiency instead of its sub-

functions. We find no evidence for multicollinearity. Variance inflation factors are all well below the threshold of 10 (Belsley 1991) and results stay similar when we estimate models that only include risk reduction or imagery benefit.

8.2 Overall Effects

Marginal means of the coefficients are $\overline{\beta_{0i,\,Germany}}$ = 0.20 and $\overline{\beta_{0i,\,USA}}$ = -0.10 for the intercept, $\overline{\beta_{1i,\,Germany}}$ = 0.74 and $\overline{\beta_{1i,\,USA}}$ = 0.85 for risk reduction, and $\overline{\beta_{2i,\,Germany}}$ = 0.18 and $\overline{\beta_{2i,\,USA}}$ = 0.14 for imagery benefit. Overall, results are comparable for both countries and support our conceptualization. Both brand functions have a positive effect on brand relevance. But the impact of the risk reduction function is far greater than the impact of the imagery benefit function. High rationality in B2B markets seems to limit the importance of imagery components.

Visibility of B2B-branded products to the end customer leads to higher brand relevance in Germany and the US ($\gamma_{1,Germany}$ = 0.03, p < 0.001; $\gamma_{1,USA}$ = 0.04, p < 0.001). Provider heterogeneity is only significant in the German model ($\gamma_{2,Germany}$ = 0.02, p < 0.01). At least for Germany, both seem to directly increase the importance of brands for B2B provider choices.

For the moderator variables, we obtain similar effects for the buying center size in Germany and the US. A larger buying center leads to a higher impact of the risk reduction function ($\lambda_{14,Germany}$ = 0.06, p < 0.01; $\lambda_{14,USA}$ = 0.19, p <0.01), while it decreases the effect of imagery benefits ($\lambda_{24,Germany}$ = -0.07, p < 0.01; $\lambda_{24,USA}$ = -0.15, p < 0.01). Individual buying center members vary in their personal backgrounds (de Ruyter, de Jong, and Wetzels 2009). Thus, the larger the buying center, the larger the variety of individual backgrounds. As evaluations of product characteristics and their contribution to purchase risk vary across members, the impact of the risk reduction function increases in the buying center size. Here, brands help

combine the different evaluations of buying center members to an overall evaluation. They work as an anchor that bundles divergent views (Spiro 1983). Contrarily, the multitude of personal backgrounds leads to a higher variety in interpretations of a brand. As brands can only contribute to an individual's self-concept if they have a unique interpretation, the imagery benefit of a brand decreases with the buying center size.

8.3 Differences between Germany and the US

We obtain several differences in moderators between Germany and the US. The strongest difference occurs for the respondents' responsibility. All three moderations show opposite directions for Germany and the US. The direct effect is positive in Germany ($\lambda_{01,\text{Germany}} = 0.44$, $p < 0.01$) and negative in the US ($\lambda_{01,\text{USA}} = -0.47$, $p < 0.01$), the impact on the effect of risk reduction is negative in Germany ($\lambda_{11,\text{Germany}} = -0.24$, $p < 0.01$) and positive in the US ($\lambda_{11,\text{USA}} = 0.37$, $p < 0.01$), and the impact on the effect of the imagery benefit is positive in Germany ($\lambda_{21,\text{Germany}} = 0.13$, $p < 0.01$) and negative in the US ($\lambda_{21,\text{USA}} = -0.25$, $p < 0.01$). In general, when taking actions alone, managers cannot rely on colleagues. Information overload and, thus, brand relevance increase. Moreover, the personal risk cannot spread among several persons and companies may more easily see managers' contribution to a decision or recommendation so that the need for risk reduction and imagery benefits increases. All effects should, hence, be positive.

We explain negative effects by the different cultures. US business culture involves a low long-term orientation (Hofstede 2016), which promotes variable payment systems based on short-term financial outcomes. Here for the individual manager, financial outcomes gain ample importance in a purchase decision. Likewise, brands have a lower relevance. In contrast, lower humane orientation in leadership (Dorfman, Hanges, and Brodbeck 2004) and high uncertainty avoidance (Hofstede 2016) shape German business culture. Here, managers may

become reluctant to rely on others' risk evaluations. The perceived risk and the need for risk reduction in group actions increases. Finally, strong B2B brands represent long-term orientation (Lennartz et al. 2015). As US-Americans value long-term orientation less, other aspects than brands gain importance for representation issues. The impact of imagery benefits decreases.

For the number of reporting Levels to the CEO, we see a slight difference in the impact of risk reduction between Germany ($\lambda_{12,Germany}$ = 0.03, p < 0.05) and the US ($\lambda_{12,USA}$ = -0.05, p < 0.10). The need for risk reduction may generally increase for top-level managers because, through their strategic focus, they burden a higher organizational risk. But German employees experience lower support from their leaders because leadership is less humane oriented (Dorfman, Hanges, and Brodbeck 2004). Here, a higher personal risk for low-level managers seems to exceed their lower organizational risk.

For larger companies, the impact of a bad purchase decision is relatively lower for the overall company performance than for smaller companies. Thus for the size of a company, a negative interaction with the risk reduction function should result. Our German model shows this negative interaction ($\lambda_{13,Germany}$ = -1.68×10^{-3}, p < 0.05), while the effect is non-significant and slightly positive in the US ($\lambda_{13,USA}$ = 0.34×10^{-3}, p > 0.10). The latter may again result from common variable payments in larger US firms. The individual risk of a bad purchase decision may increase and cancel out the lower risk for the company performance.

For the global reach of a company, local orientation reduces the influence of the risk reduction in Germany ($\lambda_{15,Germany}$ = -0.03, p < 0.10) and enhances it in the US ($\lambda_{15,USA}$ = 0.09, p < 0.01). We explain the effect for the US by the fact that local companies supply less customers than global companies. Product mistakes may, hence, affect a larger percentage of customers than for globalized companies. The risk in case of a product failure and the impact of risk reduction increases. The opposite effect in Germany may result from the structure of econom-

ic sectors. Here, industry is far more important than in the US. Industrial goods demand for joined planning processes along the supply chain. Especially due to their high interconnectedness, local businesses may choose providers in a joined decision with downstream members of the supply chain.

9 Discussion

Our study contributes to the research on B2B brands. In line with Homburg, Klarmann, and Schmitt (2010), we can show that brands actually bear relevance in B2B markets. Nevertheless, the structure of its antecedents differs to B2C markets. While social demonstration and risk reduction drive brand relevance in B2C markets, we can show that B2B customers use brands to efficiently process information. This result is in line with the finding of Zablah, Brown, and Donthu (2010) that brand preference relates stronger to brand sensitivity in highly competitive environments. Competitive environments go along with higher information load. Here, consumers' are more likely to form choice sets based on preferred brands to reduce this overload. They gain sensitivity for their preferred brands.

Furthermore, we contribute to the ongoing debate on managers' rationality. We show that, when managers make purchase decisions, they do not solely rely on hard data, but also consider brands for their evaluations. Hence, our findings add to the results of other researchers who have shown that managers take into account non-rational arguments for decisions (Davis-Sramek et al. 2009).

Information efficiency itself evolves as a reflective second-order construct with the subfunctions risk reduction and imagery benefits. Brands, thus, seem to fulfill two major tasks. They bundle information on purchase risks and image components of a brand. Contrary to existing studies (e.g., Homburg, Klarmann, and Schmitt 2010), we can show that risk reduction and information efficiency are no separable processes, but that the latter is a reflective subfunction of the former.

Importantly, while existing studies leave out imagery components of brands (e.g., Zablah, Brown, and Donthu 2010), we show that imagery benefits of B2B brands play a significant role. This result matches the finding of Lennartz et al. (2015) that intangible brand associations actually shape B2B brand strength. However, the effect of imagery benefits on brand relevance is much smaller than the effect of risk reduction. This result is in line with the results of Fischer, Voelckner, and Sattler (2010) who find for B2C markets that image-related functions of a brand are less important for brand relevance than its risk reduction potential. Moreover, our results show that it is important to consider individual heterogeneity between respondents. Although the marginal mean of the coefficients of imagery benefit is positive in both countries, coefficients turn out to be negative for 3.57% of the German and 19.30% of the US respondents when we include all interaction terms. Depending on individual characteristics, some managers seem to perceive using imagery benefits of brands as counterproductive.

Distinct to the brand functions, brand relevance increases in the visibility of a B2B branded product to the end customer in both countries. This finding matches the idea of ingredient branding. End customers do not only value brands of the end product, but also of its components. Here, visibility of components increases potential spillover benefits for the brand of the end product (Simonin and Ruth 1998). This result also supports the finding of Homburg, Klarmann, and Schmitt (2010) that B2B brand equity increases in brand awareness.

9.1 Moderators

Similar to B2C markets (Fischer, Voelckner, and Sattler 2010), respondents' personal characteristics moderate the impact of brand functions on brand relevance. In B2C markets, these characteristics are of demographic nature. Contrarily, we can show that moderators in B2B markets cover manager and organizational characteristics. Hence, they do not only differ from B2C markets, but also involve a personal and company component.

139

Not only moderators, but also the implications of their effects differ from B2C markets. While Fischer, Voelckner, and Sattler (2010) find that brand relevance is higher in group than single decisions, we find that, for the US, brand relevance generally decreases in the buying center size. This finding is in line with the finding of Homburg, Klarmann, and Schmitt (2010) that buying center size reduces the effect of B2B brand awareness on market performance.

Finally, we show for the US that with a higher degree of localization imagery benefits have lower importance. So at least in the US, customers that operate globally especially value image-related associations of brands. This result is in line with Torres et al. (2012) who show that corporate social responsibility as an imagery component is a driver of global brand equity.

9.2 Differences between Countries

Interestingly, variation in brand relevance between categories is substantially higher in Germany than the US. This supports the notion of Fischer, Voelckner, and Sattler (2010) that brand relevance does not exist per se, but companies may build it by investments into brands. While B2B companies in Germany in some categories have been highly restricted and only recently started to invest into brands, companies in the US have discovered investments into B2B brands for a longer time.

Individual interactions of moderator variables partly show strong differences between Germany and the US. The most striking difference in moderators exists for respondents' responsibility for their actions. We show that the direct effect of acting alone on brand relevance is positive in Germany and negative in the US. This result matches the finding of Lennartz et al. (2015). They find for the US that, if managers act alone, the brand strength of a company's primary provider in a category is lower than, if managers act in a group, while no

effect exists for Germany. So individual brand perception mechanisms seem to differ between both countries.

Overall, moderation effects reflect that German business culture features a higher long-term orientation and uncertainty avoidance (Hofstede 2016), but a lower humane orientation (Dorfman, Hanges, and Brodbeck 2004). Our results, thus, support the view of Dorfman, Hanges, and Brodbeck (2004) and Ronen and Shenkar (2013) that Germany and the US belong to different clusters of business culture. Altogether, these notions help answer the question of Keller and Lehmann (2006) how perceptions of branding strategies vary across countries.

10 Management Implications

Our research contributes to managerial decision making in several ways. First, we provide companies with a tool to measure B2B brand relevance. As we validate this tool across two countries and seven generic B2B categories, application should be meaningful across other countries and categories. Companies may use this variation to forecast the profitability of B2B brand building and, thus, the allocation of marketing budget across countries and categories.

Second, we find that brands are relevant in B2B markets across seven product categories and two countries. While product categories in a country reveal only slight differences, differences are larger between countries. It seems to be a valid strategy for international B2B companies to allocate their brand-building budget across countries. Nevertheless, for a more detailed categorization, brand relevance may show stronger variation across categories.

Third, we find both the risk reduction function and the imagery benefit function to drive brand relevance across countries. Herein, the effect of imagery benefits is smaller than the effect of risk reduction. It, hence, seems to be a valid strategy to especially strengthen risk

reduction potential of brands in B2B branding strategies. Nevertheless, although managers may less pronounce image components, neglecting them may lead to unfavorable results.

Fourth, we analyze the influence of manager and organizational characteristics on brand relevance. As B2B purchases involve intense relationships, companies should train their sales force to use these individual influences to develop tailored communication strategies. For both countries, we see that a larger buying center size increases the importance of the risk reduction function, but decreases the importance of the imagery benefit function. Companies may thus emphasize risk reduction elements of their brand for large buying centers, but deemphasize imagery elements. Finally, other moderations, e.g., for a respondents' responsibility, vary between Germany and the US. Further country-specific distinctions seem fruitful in brand communication.

11 Conclusion and Avenues for Further Research

We close an important research gap in B2B marketing. First, against ongoing doubts, we can show that B2B brands actually bear relevance. We develop and test a conceptual framework that shows how B2B brand relevance arises. By showing that imagery benefits generate brand relevance, we overcome an important limitation of former research. Our findings particularly strengthen the view that researchers need to consider B2B brands when they study company interactions. However, it is not sufficient to focus on brands' risk reduction potential. They also need to include B2B brands' imagery component because these components may substantially add to the relevance and, thus, the impact of B2B brands for purchase decisions.

Second, we introduce a scale that measures B2B brand relevance. As we validate it across two countries and seven product categories, researchers can use it to assess B2B brand relevance as a construct in various future research settings. Including this construct, helps understand in which situations "soft" arguments, such as brands, promote sales.

142

Third, while previous studies have focused only on one category or country, we can show that B2B brand relevance and functions vary across countries and categories. In this context, brands add to the understanding of cross-category and -country differences in B2B purchase processes.

Fourth, we show that manager and organizational characteristics lead to further variation in B2B brand relevance. We can show that these characteristics have ample implications for B2B brand management. In particular, they strengthen different functions of brands. Thus, if researchers study B2B brand strategies, it is not sufficient to rely on a company or business perspective, but they need to consider the alignment of these strategies to the individual manager. Otherwise, they may over- or underestimate the influence of the specific brand functions and associations.

Our study is prone to few limitations. First, although we cover two countries with large economic relevance, the influences of brand functions and their moderation by manager and organizational characteristics may differ in other countries. Especially, developing countries may reveal promising results for future research. Second, we do not include actual market outcomes in our study. Future research should assess how B2B brand relevance combined with the actual brand strength influence market or financial outcomes. Finally, our study only involves large firms with more than 100 employees. Future research should assess whether B2B brand relevance and processes differ for smaller companies. Nevertheless, we expect branding issues to be less relevant for smaller companies and, thus, our sample to be correspondingly accurate.

REFERENCES PROJECT 3

Aaker, David A. and Robert Jacobson, (2001). "The Value Relevance of Brand Attitude in High-technology Markets," *Journal of Marketing Research*, 38 (4), 485-493.

Atkinson, Richard C. and Richard M. Shiffrin (1968). "Human Memory: A Proposed System and Its Control Processes," in K. W. Spence, J. T. Spence (Eds.), *The Psychology of Learning and Motivation*, New York: Academic Press, 89-195.

Bagozzi, Richard P. and Youjae Yi (1988). "On the Evaluation of Structural Equation Models," *Journal of the Academy of Marketing Science*, 16 (1), 74-94.

Belsley, David A. (1991). *Conditioning Diagnostics: Collinearity and Weak Data Regression*, New York: Wiley.

Bollen, Kenneth A. (1989). *Structural Equations with Latent Variables*, New York et al.: Wiley.

Brown, Tom J. and Peter A. Dacin (1997). "The Company and the Product: Corporate Associations and Consumer Product Responses," *Journal of Marketing*, 61 (1), 68-84.

Brown, Brian P., Alex R. Zablah, Danny N. Bellenger, and Wesley J. Johnson (2011). "When do B2B Brands Influence the Decision Making of Organizational Buyers? An Examination of the Relationship between Purchase Risk and Brand Sensitivity," *International Journal of Research in Marketing*, 28 (3), 194-204.

Bureau of Economic Analysis (2016). *Gross-domestic-product-by-industry Data* [available at http://www.bea.gov/industry/gdpbyind_data.htm; accessed March 3, 2016].

Cleeren, Kathleen, Harald J. van Heerde, and Marnik Dekimpe (2013). "Rising from the Ashes: How Brands and Categories Can Overcome Product-harm Crises," *Journal of Marketing*, 77 (2), 58-77.

Cooper, Marjorie J., Kirk L. Wakefield, and John F. Tanner (2006). "Industrial Buyers' Risk Aversion and Channel Selection," *Journal of Business Research*, 59 (6), 653-661.

Darby, Michael R. and Edi Karni (1973). "Free Competition and the Optimal Amount of Fraud," *Journal of Law and Economics*, 16 (1), 67-88.

Davis-Sramek, Beth, Cornelia Droge, John T. Mentzer, and Matthew B. Myers (2009). "Creating Commitment and Loyalty Behavior among Retailers: What Are the Roles of Service Quality and Satisfaction?," *Journal of the Academy of Marketing Science*, 37 (4), 440-454.

De Ruyter, Ko, Ad de Jong, and Martin Wetzels (2009). "Antecedents and Consequences of Environmental Stewardship in Boundary-spanning B2B Teams," *Journal of the Academy of Marketing Science*, 37 (4), 470-487.

Desai, Kalpesh K., and Kevin L. Keller (2002). "The Effects of Ingredient Branding Strategies on Host Brand Extendibility," *Journal of Marketing*, 66 (1), 73-93.

DelVecchio, Devon, Cheryl B. Jarvis, Richard R. Klink, and Brian R. Dineen (2007). "Leveraging Brand Equity to Attract Human Capital," *Marketing Letters*, 18 (3), 149-164.

Dorfman, Peter W., Paul J. Hanges, and Felix C. Brodbeck (2004). "Leadership and Cultural Variation: The Identification of Culturally Endorsed Leadership Profiles," in R. J. House, P. J. Hanges, M. Javidan, P. W. Dorfman, V. Gupta (Eds.), *Culture, Leadership, and Organizations*, Thousand Oaks, CA: Sage, 669-720.

Fischer, Marc, Wolfgang Giehl, and Tjark Freundt, T. (2011). "Managing Global Brand Investments at DHL," *Interfaces*, 41 (1), 35-50.

Fischer, Marc, Franziska Völckner and Henrik Sattler (2010). "How Important Are Brands? A Cross-category, Cross-country Study," *Journal of Marketing Research*, 47 (5), 823-839.

Ghosh, Mrinal and George John (2009). "When Should Original Equipment Manufacturers Use Branded Component Contracts with Suppliers?," *Journal of Marketing Research*, 46 (5), 597-611.

Grubb, Edward L. and Harrison L. Grathwohl (1967). "Consumer Self-concept, Symbolism, and Market Behavior: A Theoretical Approach," *Journal of Marketing*, 31 (4), 22-27.

Grubb, Edward L. and Gregg Hupp (1968). "Perception of Self, Generalized Stereotypes, and Brand Selection," *Journal of Marketing Research*, 5 (1), 58-63.

Hawes, Jon M. and Scott H. Barnhouse (1987). "How Purchasing Agents Handle Personal Risk," *Industrial Marketing Management*, 16 (4), 287-293.

Hofstede, Geert (2016). *Cultural Dimensions* [available at http://geert-hofstede.com/countries.html, accessed November 13, 2016].

Homburg, Christian., Martin Klarmann, Martin Reimann, and Oliver Schilke (2012). "What Drives Key Informant Accuracy?," *Journal of Marketing Research*, 49 (4), 594-608.

Homburg, Christian., Martin Klarmann, and Jens Schmitt (2010). "Brand Awareness in Business Markets: When Is It Related to Firm Performance?," *International Journal of Research in Marketing*, 27 (3), 201-212.

Hughes, D. E., & Ahearne, M. (2010). "Energizing the Reseller's Sales Force: The Power of Brand Identification," *Journal of Marketing*, 74 (1), 81-96.

Johnston, Wesley L. and Jeffrey E. Lewin (1996). "Organizational Buying Behavior: An Integrative Framework," *Journal of Business Research*, 35 (1), 1-15.

Keller, Kevin L. (1993). "Conceptualizing, Measuring, and Managing Customer-based Brand Equity," *Journal of Marketing*, 57 (1), 1-22.

Keller, Kevin L. and Donald R. Lehmann (2006). "Brands and Branding: Research Findings and Future Priorities," *Marketing Science*, 25 (6), 740-759.

La, Vinh, Paul Patterson, and Chris Styles (2009). "Client-perceived Performance and Value in Professional B2B Services: An International Perspective," *Journal of International Business Studies*, 40 (2), 274-300.

Lemke, Fred, Moira Clark, and Hugh Wilson (2011). "Customer Experience Quality: An Exploration in Business and Consumer Contexts Using Repertory Grid Technique," *Journal of the Academy of Marketing Science*, 39 (6), 846-869.

Lennartz, Eric M., Marc Fischer, Manfred Krafft, and Kay Peters (2015). "Drivers of B2B brand strength – Insights from an international study across industries," *Schmalenbach Business Review*, 67 (1), 114-137.

Löhndorf, Birgit, and Adamantios Diamantopoulos (2014). "Internal branding: social identity and social exchange perspectives on tuning employees into brand champions," *Journal of Service Research*, 17 (3), 310-325.

Luo, Anita and V. Kumar (2013). "Recovering Hidden Buyer-seller Relationship States to Measure the Return on Marketing Investments in Business-to-business Markets," *Journal of Marketing Research*, 50 (1), 143-160.

Miller, George A. (1956). "The Magical Number of Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information," *Psychological Review*, 63 (2), 81-97.

Nelson, Phillip (1970). "Information and consumer behavior," *Journal of Political Economy*, 78 (2), 311-329.

Netemeyer, Richard G., William O. Bearden, and Subash Sharma (2003). *Scaling Procedures*. Thousand Oaks et al.: Sage.

Pick, Doreén and Martin Eisend (2014). "Buyers' Perceived Switching Costs and Switching: A Meta-analytic Assessment of Their Antecedents," *Journal of the Academy of Marketing Science*, 42 (4), 186-204.

Ronen, Simcha and Oded Shenkar (2013). "Mapping World Cultures: Cluster Formation, Sources and Implications," *International Journal of Business Studies*, 44 (9), 867-897.

Simonin, Bernard L. and Julie A. Ruth (1998). "Is a Company Known by the Company It Keeps? Assessing the Spillover Effects of Brand Alliances on Consumer Brand Attitudes," *Journal of Marketing Research*, 35 (1), 30-42.

Smith, Peter B., Mark F. Peterson and Shalom H. Schwartz (2002). "Cultural Values, Sources of Guidance, and Their Relevance to Managerial Behavior," A 47 Nation Study," *Journal Cross-Cultural Psychology*, 33 (2), 188-208.

Spiro, Rosann L. (1983). "Persuasion in Family Decision-making," *Journal of Consumer Research*, 9 (4), 393-402.

Srinivasan, V. and Chan S. Park (1997). "Surprising Robustness of the Self-Explicated Approach to Customer Preference Structure," *Journal of Marketing Research*, 34 (2), 286-291.

Stahl, Florian, Mark Heitmann, Donald R. Lehmann, and Scott A. Neslin (2012). "The Impact of Brand Equity on Customer Acquisition, Retention, and Profit Margin," *Journal of Marketing*, 76 (4), 44-63.

Standard & Poor's (2016). *Global Industry Classification Standard (GICS) Methodology* [available at https://us.spindices.com/resource-center/index-policies/; accessed November 11, 2016].

Steenkamp, Jan-Benedict E. M. and Hans Baumgartner (1998). "Assessing Measurement Invariance in Cross-national Consumer Research," *Journal of Consumer Research*, 25 (1), 78-90.

Strategy Analytics (2016). *United States Top Ad Spending per Capita in 2014 at US$ 567 per Person* [available at https://www.strategyanalytics.com/; accessed September 6, 2016].

Torres, Anna, Tammo H. Bijmolt, Josep A. Tribó, and Peter Verhoef (2012). "Generating Global Brand Equity through Corporate Social Responsibility to Key Stakeholders," *International Journal of Research in Marketing*, 29 (1), 13-24.

World Bank. 2016. *Gross Domestic Product Ranking Table [available at* http://databank.worldbank.org; accessed March 3, 2016].

Wuyts, Stefan, Peter C. Verhoef, and Remco Prins (2009). "Partner Selection in B2B Information Service Markets," *International Journal of Research in Marketing*, 26 (1), 41-51.

Yoon, Yeosun, Zeynep Gürhan-Canli, and Beyza Bozok (2006). "Drawing Inferences about Others on the Basis of Corporate Associations," *Journal of the Academy of Marketing Science*, 34 (2), 167-173.

Zablah, Alex R., Brian P. Brown, and Naveen Donthu (2010). "The Relative Importance of Brands in Modified Rebuy Purchase Situations," *International Journal of Research in Marketing*, 27 (3), 248-260.

APPENDIX PROJECT 3

Appendix A: Items for the Validity Measures

Table A1: Items for the Validity Measures

Scale	Items
Price premium[1]	I prefer to buy strong brands even at a higher price.
Brand consideration[1]	Primarily, I check the offers of my preferred brand and then check other options.
Brand strength[1][2]	Please rate *(provider)*[3] as your current provider of *(category)*[3] with regard to the following criteria. A company 1. … with a good reputation. 2. … I trust. 3. … I feel positive about. 4. … I know well. 5. … that is clear on what it stands for.
Brand importance	Please tell us how important the following factors are for you and your company when choosing which provider to purchase *(category)*[3] from. Please assign a total of 100 points among the following factors. The more points you assign the more influence the factor has on the decision: • Brand-related factors, • Channel-related factors, • Price-related factors, • Product-related factors, • Communication-related factors.

1) we measure items on a fully labeled Likert scale (1 = strongly disagree, 2 = disagree, 3 = neither agree nor disagree, 4 = agree, 5 = strongly agree); 2) respondents could choose three out of five items; 3) we replace *(provider)* by the provider name and *(category)* by one of our seven product categories.

PROJECT 4: CULTURE AND SURVEY RESEARCH: A REVIEW ON HOW RESPONSE STYLES DIFFER ACROSS CULTURES

Authors: Eric Michael Lennartz

ABSTRACT

Globalization has fostered the mixing of cultures and the internationalization of marketing issues. Accordingly in survey research, the number of cross-cultural studies increases and even samples gathered within one country include respondents with multiple cultural backgrounds. This cultural diversity may lead to differences in response styles which contaminate scale values and their correlations. If researchers do not consider these response styles, they may draw wrong conclusions. Relevant literature is vast, but lacks structure and generalizations. Studies involve various cultures, come from multiple disciplines, and differ in their research approaches. Thus, this paper reviews current literature on how culture may shape response styles. Based on this review, it develops a framework for the effect of culture on response styles. In addition, it derives empirical generalizations for differences in response styles across cultures. Both, the framework and the generalizations, provide intuition for researchers and practitioners on how to consider potential contamination from cross-cultural surveys and samples.

1 Introduction

Challenges for marketers today differ from those years ago. While they previously have run regional or national campaigns, traditional boarders blur and marketing actions face a multinational environment in a global marketplace (Strizhakova and Coulter 2013). Thus, market researchers have to assess success factors of marketing campaigns in an increasing culturally divers environment. Researchers have reacted to these movements and conduct cross-cultural studies. The body of these cross-cultural studies is constantly increasing. Not least, Psychology and Marketing (2016) particularly invites cross-cultural studies in its Aims and Scope and, has published several cross-cultural studies in 2016 (e.g., Buzova, Sanz-Blas, and Cervera-Taulet 2016; Kim and Yi 2016).

Furthermore, cultural diversity does not only increase through internationalization, but populations within countries become more culturally diverse. For instance, in 2015, more than 20% of the German population had a migration background (Statistisches Bundesamt, 2016). Other states show even higher numbers and percentages are going to increase due to globalization, ongoing migration, and refugee flows. Likewise, migrants' cultural backgrounds differ from the cultural makeup of their host country. In an acculturation process, they learn about and decide how to adopt to the new culture (Ownbey and Horridge 1997) and apply strategies to avoid conflicts with the indigenous culture (e.g., Peñaloza 1994). For the immigrant, this process leads to new cultural patterns, which may even incorporate elements new to his/her home and the new host culture (e.g., Wallendorf and Reily 1983). Cultural diversity and pluralism are, hence, going to increase.

In this context, both, cross-cultural studies and national studies with immigrants and indigenes, bear substantive risks for survey research. In particular, a respondent's cultural background can shape response styles in consumers' answers to survey items (e.g., Steenkamp, de Jong, and Baumgartner, 2010). These response styles are tendencies "to respond systematical-

ly to questionnaire items on a basis other than what the items were specifically designed to measure" (Paulhus, 1991, p. 17). For example, respondents may show a midpoint response style. In this case, they excessively choose the midpoint category across survey items regardless of the item content (Baumgartner and Steenkamp 2001). Other response styles involve the overuse of extreme or socially desirable answer categories (Greenleaf 1992; de Jong, Pieters, and Fox 2010).

Importantly, response styles can contaminate construct values and their correlations (Steenkamp, de Jong, and Baumgartner 2010). Thus, researchers need to consider potential cultural differences in response styles. Otherwise, they may draw wrong conclusions, if they find differences between cultural groups in survey data (Schwarz 2003). They may assign these differences to cultural influences, although they simply represent differences in response styles. Likewise, this issue is of particular importance for practitioners. Survey research is and has been the major tool to discover latent constructs, such as brand experience (e.g., Brakus, Schmitt, and Zarantello 2009) and price consciousness (e.g., Lichtenstein, Ridgway, and Netemeyer 1993). These constructs offer essential insights for companies. Biases due to culture and acculturation processes may, hence, endanger companies' businesses.

While marketing literature on these issues is sparse, extensive studies from related disciplines test the effects of culture on response styles (e.g., Green 1996; Lalwani, Shavitt, and Johnson 2006). This evidence is scattered, often contradictory, and lacks generalizations. Importantly, to the best of the authors' knowledge, no framework exists which allows classifying the effects of culture on response styles. Such a framework would help marketing researchers and practitioners to systematically study and consider the effects of response styles on answers to survey items. Moreover, contradictory results and lacking generalizations lead to inconsistent recommendations. For example, some studies recommend to use a higher number of answer categories for an item to avoid differences in extreme response styles between cul-

tures (e.g. Hui and Triandis 1989), while others recommend less answer categories for the same reason (e.g., Clarke 2000).

Thus, the objective of this study is a comprehensive literature review of studies that assess how culture may shape response styles. It intends to offer guidance for marketing researchers and practitioners on the interplay of culture and response styles. Therefore, this study first introduces a framework on the general influences of culture on response styles to questionnaire items and potential moderators. Researchers and marketers can use this framework to get an intuition on potential factors that they should consider when they conduct studies across culturally diverse samples. Based on this framework, the study classifies, analyzes, and sums up general results from empirical studies. Researchers and managers can use these generalizations to evaluate whether culture may contaminate their results. In this context, the study also identifies research gaps that researchers can take as avenues for further research.

The structure of the remainder of the paper is as follows: The following section introduces a framework that covers the influences of culture on the formation of response styles. The study proceeds with a section that assesses empirical studies as classified by the given framework. Based on these findings, it illustrates existing research gaps which are avenues for further research. The study closes with management implications and a conclusion.

2 A Conceptual Framework for the Influence of Culture on Response Style

Figure 1 displays the overall framework of the literature review of this study. If respondents participate in a survey, they usually answer several items that form a questionnaire. These items are statements or questions that respondents have to answer in a given format (e.g., Tourrangeau, Rips, and Rasinski 2009). They consist of content-related and measurement characteristics. The latter split into factors such as language, item polarity, scale format, number of answer categories, and labeling of answer categories.

Figure 1: Conceptual Framework for the Effect of Culture on Response Styles

Questionnaire Items	Response Process	Response Styles

Questionnaire Items
- content-dependent characteristics
- language
- item polarity
- scale format
- number of answer categories
- labeling of answer categories

Response Process
- Comprehension
- Retrieval
- Judgment
- Response

Response Styles
- content-dependent
- content-independent

Personal Characteristics

Culture
- beliefs, values, and norms
- acculturation processes
- global culture

When respondents answer questionnaire items they follow four steps, namely comprehension, retrieval, judgment, and response (Tourangeau, Rips, and Rasinski 2009). In all steps, errors may occur so that respondents deviate in their observed answer score from their true answer scores (Johnson et al. 1997). These errors depend on respondents' personal characteristics. They determine the basic principles for all steps of the response process (Tourangeau, Rips, and Rasinski 2009). Among these characteristics, culture provides the beliefs, values, and norms that respondents use for their item response (e.g., Johnson et al. 1997; Lau-Gesk 2003).

If a deviation between observed and true answer score occurs nonrandomly, it represents a response style (Baumgartner and Steenkamp 2001). Notably, response styles can occur independent of respondents' intention to show such style. They may, thus, occur consciously and nonconsciously.

Importantly, response styles contribute to systematic measurement error. They may contaminate not only the observed item scores, but also the relationships between these values (Baumgartner and Steenkamp 2001). Therefore, response styles may inflate or deflate both, scale values and correlations between scale values, and, thus, harm the results of further statistical analyses.

2.1 Types of Response Styles

Table 1 provides an overview on major response styles and their definitions. Overall, they split up into two groups[3]. A first group is content-independent. Here, response styles predominantly show as a pattern in ticking specific answer categories, such as midpoints (e.g., Baumgartner and Steenkamp 2001). In addition, such styles can result from careless and overly fast response or situations where respondents give no answers.

Table 1: Major Response Styles and Definitions

Response Style		Definition
Content-independent response styles		
ARS	Acquiescence Response Style	"Tendency to agree with questionnaire or personality items regardless of content" (Bachman and O'Malley 1984, 491).
DARS	Disacquiescence Response Style	"Tendency to disagree with items regardless of content" (Baumgartner and Steenkamp 2001, 145).
NARS	Net Acquiescence Response Style	"Tendency to show greater acquiescence than disacquiescence" (Baumgartner and Steenkamp 2001, 145)
ERS	Extreme Response Style	"Tendency to favor […] using the endpoints of a rating scale relatively independent of specific item content" (de Jong et al. 2008, 104).
MRS	Midpoint Response Style	"Tendency to use the middle scale category regardless of content" (Baumgartner and Steenkamp 2001, 145).
RR	Response Range	"Tendency to use the entire range of the scale in responding" (Hui and Triandis 1989, 303).[1)]

---

[3]Some authors refer to the first group as response styles and the second group as response sets and combine both under the term response bias (e.g., Watkins and Cheung 1995). However, as most studies use these terms interchangeably (e.g., Baumgartner and Steenkamp 2001), this study refers to all relevant phenomena as response styles.

Table 1: Major Response Styles and Definitions (Cont'd)

Response Style		Definition
Content-independent response styles		
NCR	Noncontingent Response Style	Tendency to apply a "random pattern that does not take the content of the items into account" (Watkins and Cheung 1995, 494) and "to respond carelessly, randomly, and nonpurposefully" (Baumgartner and Steenkamp 2001, 145).
IO	Item Omission	Tendency to state no response to items regardless of their content (Webster 1996).
S	Speed	"Tendency to work quickly" (Broen and Wirt 1958, 238) when answering survey items.[2]
TL	Tendency to the Left	"Tendency to rate to the left […] of center" regardless of content (Broen and Wirt 1958, 237)
TR	Tendency to the Right	"Tendency to rate to the […] right of center" regardless of content (Broen and Wirt 1958, 237)
FEB	Frequency Estimation Bias	Tendency of respondents to "use the range of values provided on the scale as a frame of reference estimating their own behavioral frequency" (Ji, Schwarz, and Nisbett 2000, 586)
Content-dependent response styles		
SDR	Social Desirable Responding	Tendency to choose answers "that make the respondent look good, based on cultural norms about the desirability of certain traits, attitudes, interests, opinions, and behavior" (Steenkamp, de Jong, and Baumgartner 2010, 200).
ERT	Egoistic Response Tendency	Tendency to show oneself "as exceptionally talented" (Paulhus and John 1998, 1034).
MRT	Moralistic Response Tendency	Tendency to show oneself "as exceptionally good member of society" (Paulhus and John 1998, 1034).

1) some studies (e.g., Fischer, Völckner, and Sattler 2010) refer to a uniform response style. As this response style is simply the opposite of RR, this study does not consider it separately; 2) Speed itself may not necessarily lead to a response style, but may be the origin of other response styles like Noncontingent Response Styles.

Most prominent among these response styles are Acquiescence Response Styles (ARS), Disacquiescence Response Styles (DARS), Extreme Response Styles (ERS), Midpoint Response Styles (MRS), and the spread of a respondent's Response Range (RR). ARS refers to a respondent's tendency to agree, while DARS refers to the tendency to disagree on items regardless of their content (Bachman and O'Malley 1984; Baumgartner and Steenkamp 2001). ERS covers a respondent's tendency to use extreme answer categories on both sides, while

MRS indicates a respondent's tendency to excessively choose the midpoint answer category (Baumgartner and Steenkamp 2001; de Jong et al. 2008). Finally, RR deals with a respondent's tendency to spread answers across the full spectrum of answer categories in comparison to show a uniform distribution of answers (Hui and Triandis 1989).

A second group covers those response styles that occur, if respondents depict a desired picture of themselves by their answers (e.g., Steenkamp, de Jong, and Baumgartner 2010). These response styles depend on the particular content of the item. Researchers summarize them under the term Socially Desirable Responding (SDR). They split up into two distinct styles. Egoistic Response Tendencies (ERT) cover those types of SDR that enhance the perception of respondents skills concerning intellect, emotions, and social competence (Steenkamp, de Jong, and Baumgartner 2010). On the contrary, Moralistic Response Tendencies (MRT) show the respondent as more altruistic, responsible, and caring for personal interrelationships (Paulhus and John 1998; Steenkamp, de Jong, and Baumgartner 2010).

2.2 Influence of Culture on Response Styles

2.2.1 Beliefs, Values, and Norms

According to the trait concept in cross-cultural psychology, an individual's personality consists of specific manifestations of personality traits (Church 2000). These traits represent the initial ways of individuals' perception and behavior (Steenkamp, de Jong, and Baumgartner 2010). In this context, culture serves as a framework that outlines the manifestations of personality traits. It represents "the beliefs, values, and norms of a specific sociocultural group" (Luna, Ringberg, and Perracchio 2008, p. 280). When individuals form their identities, they learn these beliefs, values, and norms through interaction with their environments, for example, via personal contacts or exposure to media (Lee 1993; Peñaloza 1989).

For item response, culture influences all four steps of the response process. Comprehension may differ due to different connotations and connections of specific words across cultures (Luna, Ringberg, and Perracchio 2008). Retrieval occurs according to general schemes and depends on respondents' previous knowledge. Both factors may vary across cultures (D'Andrade 1995; Ji, Schwarz, and Nisbett 2000). Judgment involves several steps, such as agreeing with certain positions, where the general consensus across cultures may differ (e.g., Hofstede 2001; Steenkamp, de Jong, and Baumgartner 2010). Finally, certain response categories may represent different features across cultures that motivate respondents to edit their responses (e.g., Hui and Triandis 1989; de Langhe et al. 2011).

2.2.2 Acculturation Process

Importantly, sociocultural environments may change, if individuals migrate to new countries. While indigenes have internalized only one culture, immigrants face two likely different cultures, the culture of their home country and the culture of the new host country (e.g., Askegaard, Arnould, and Kjeldgaard 2005). Immigrants are aware of differences between these cultures and, hence, apply strategies to cope with them (Peñaloza 1994). They become encompassed in an acculturation process. Here, they develop new manifestations of personality traits, identities and practices that allow more convenient interaction in the new cultural environment (Lee 1993; Luedicke 2015; Ownbey and Horridge 1997). During this process, they take over elements of the new host culture, preserve elements of their home culture, and involve adaptive elements (Peñaloza 1989). It can lead to assimilation, integration, separation, marginalization or similar modes (Berry 1986; Üstüner and Holt 2007)[4].

---

[4] Assimilation describes the pattern when immigrants fully take over the indigenes' culture. Integration means the parallel existence of home and host culture. Immigrants maintain their home culture, but often interact with individuals from their host culture. Separation is the pattern where immigrants avoid interaction with the host culture and preserve their initial culture. Finally, marginalization results, when individuals desire to assimilate to the new culture, but miss the abilities due to low levels of interactions with individuals from the host culture.

Traditionally, adaptive elements are those components of an individual's culture that are hybrid between the home and host cultures (e.g., Faber, O'Guinn, and McCarty 1987). On the contrary, studies show that other forms exist. For example, individuals might have an image of the host culture and "hyperassimilate" to this image (Wallendorf and Reily 1983). In these cases, they show expected patterns of the host culture in even more extreme ways than indigenes (e.g., Goldsmith, White, and Stith 1987). Importantly, immigrants may not necessarily only possess one culture. They may even incorporate a full set of cultural beliefs, values, and norms from their home and host culture (Luna, Ringberg, and Perracchio 2008). Depending on external cues and their acculturation level, they may show a behavior that emphasizes elements of either culture, but are not restricted to one culture (e.g., Askegaard, Arnould, and Kjeldgaard 2005; Chattaraman, Lennon, and Rudd 2010). Thus, acculturation may induce additional cultural patterns that shape response styles.

2.2.3 Global Culture

Globalization induces additional forces that may challenge existing cultural frameworks. Global consumption trends are present in media (e.g., Kjeldgaard and Askegaard 2006), numerous multinational corporations have already developed (Strizhakova and Coulter 2013), and global brands spread across countries (e.g., Torres et al. 2012). Subsequently, general cultural trends emerge across formerly distinct national cultures that sum up to a global culture. While traditional approaches only divide cultures along ethnicity, race, or nationality, recent marketing studies consider global cultural trends and their interplay with local elements (e.g., Kjeldgaard and Askegaard 2006; Strizhakova and Coulter 2013). In detail, individuals may experience an acculturation to this global culture (Askegaard, Arnould, and Kjeldgaard 2005). This acculturation is not exclusive for immigrants, but also affects indigenes. However, receptivity to this global culture is likely to differ across traditional cultures (Kumar et al.

2013) and other personal characteristics (e.g., Strizhakova and Coulter 2013). Global culture may, hence, lead to further cultural patterns that shape response styles.

3   Empirical Findings concerning Response Styles and Culture

Subsequently, this study analyzes empirical findings concerning the impact of culture on response styles. Hereby, it focuses on the most prominent response styles, ARS, DARS, MRS, ERS, RR, SDR, ERT, and MRT and examines corresponding results. Table A1 in the Appendix gives an overview on all empirical studies reviewed.

3.1   Classification of Empirical Findings

Figure 2 shows the classification of empirical works used in this study. It divides studies into four streams.

Authors from the first stream compare response styles between countries. Most of theses studies compare respondents from Asian and Anglo[5] cultures. A second stream compares immigrants and indigenes within one country. Their focus lies on studies that compare Hispanic and African Americans to non-Hispanic white Americans. For the third stream, authors analyze response styles across multiple countries. Here, making pairwise comparisons would lead to disproportionate efforts. Thus, most of the studies compare response styles along the dimensions individualism, masculinity, power distance, and uncertainty avoidance by Hofstede (2001). Table 2 below provides definitions for these dimensions.

---

[5] The term Anglo summarizes those countries that share an English heritage, such as Australia, Canada, UK, and USA. Although they do not necessarily display geographical closeness, various studies have shown that they from a common cultural cluster due to their shared heritage (e.g., Ronen and Shenkar, 2013).
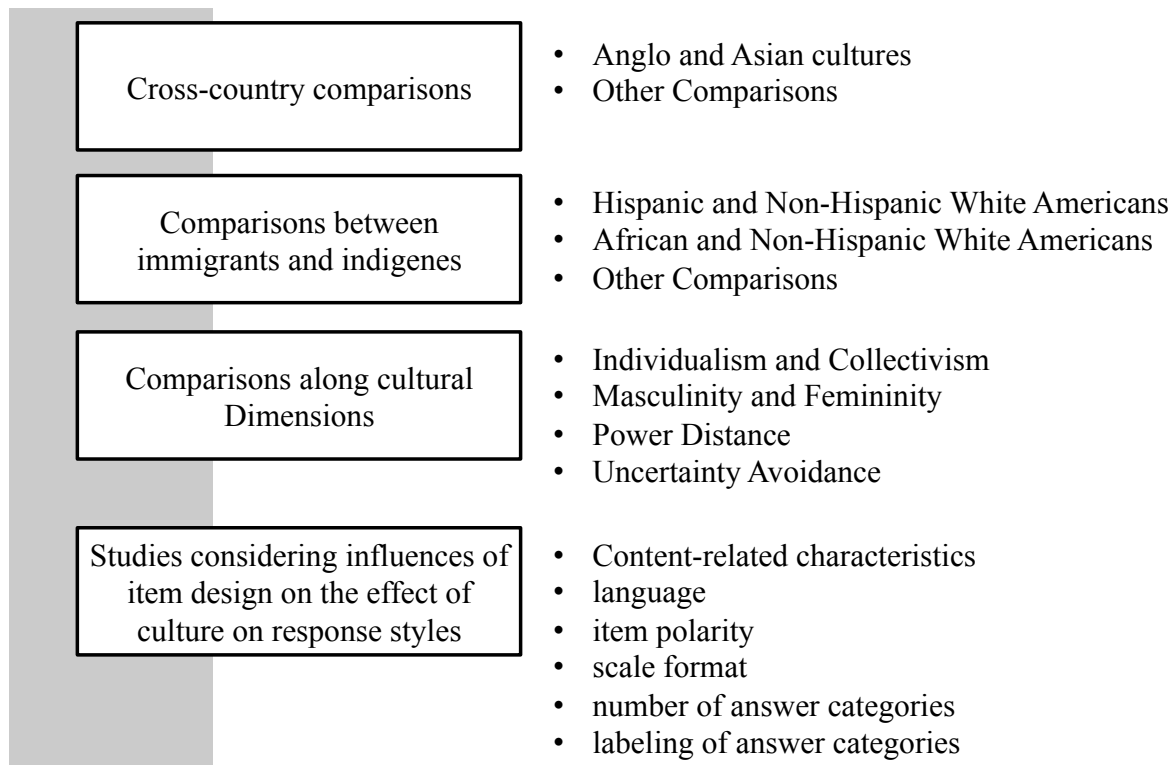
Figure 2: Classification of Empirical Studies

Cross-country comparisons	• Anglo and Asian cultures • Other Comparisons
Comparisons between immigrants and indigenes	• Hispanic and Non-Hispanic White Americans • African and Non-Hispanic White Americans • Other Comparisons
Comparisons along cultural Dimensions	• Individualism and Collectivism • Masculinity and Femininity • Power Distance • Uncertainty Avoidance
Studies considering influences of item design on the effect of culture on response styles	• Content-related characteristics • language • item polarity • scale format • number of answer categories • labeling of answer categories

Table 2: Definitions of Cultural Dimensions by Hofstede (2001)

Cultural Dimension by Hofstede (2001)	Definition
Individualism/ Collectivism	This dimension refers to "the integration of individuals into primary groups" (Hofstede 2001, p. 29). "[I]ndividualism, can be defined as a preference for a loosely-knit social framework in which individuals are expected to take care of only themselves and their immediate families. Its opposite, collectivism, represents a preference for a tightly-knit framework in society in which individuals can expect their relatives or members of a particular in-group to look after them in exchange for unquestioning loyalty" (Hofstede 2016).
Masculinity/ Femininity	This dimension refers to "the division of emotional roles between men and women" (Hofstede 2001, p. 29). "Masculinity […] represents a preference in society for achievement, heroism, assertiveness and material rewards for success" (Hofstede 2016). "[F]emininity, stands for a preference for cooperation, modesty, caring for the weak and quality of life.
Power Distance	This dimension refers to "the different solutions to the basic problem of human inequality" (Hofstede 2001, p. 29). It "expresses the degree to which the less powerful members of a society accept and expect that power is distributed unequally" (Hofstede 2016).
Uncertainty Avoidance	This dimension refers to "the level of stress in a society in the face of an unknown future" (Hofstede 2001, p. 29). It "expresses the degree to which the members of a society feel uncomfortable with uncertainty and ambiguity" (Hofstede 2016).

Finally, studies analyze how the characteristics of survey items influence the effects of culture on response styles. In line with the framework in figure 1, this review classifies results by content-related characteristics, language, item polarity, scale format, number of answer categories, and labeling of answer categories.

3.2 Cross-country Comparisons

Table 3 sums up results for the most prominent comparisons between countries and immigrants and indigenes.

Table 3: Selected Differences of Cultures in Response Style[1),2)]

	East Asians vs. Anglos	Hispanics vs. non-Hispanic white Americans	African vs. non-Hispanic white Americans
Acquiescence Response Style	(+)	(+)	(+)
Disacquiescence Response Style	(-)	(-)	(+)
Extreme Response Style	(-)	(+)	(+)
Midpoint Response Style	(+)	*no studies*	*no studies*
Response Range	(+)	*no studies*	*no studies*
Egoistic Response Tendency	(-)	(+)[3)]	(+)[3)]
Moralistic Response Tendency	(+)	(+)[3)]	(+)[3)]

1) only those comparisons of cultural groups were selected that have been analyzed in more than four studies; 2) cells show tendencies for the first culture named compared to the second culture named; 3) studies only consider social desirability on a general level, but not its components.

For cross-country comparisons, it covers results for differences between East Asian[6] and Anglo cultures. For comparisons between immigrants and indigenes, it covers differences between Hispanic and non-Hispanic white Americans and African and non-Hispanic white

---

[6] In this study East Asia refers to China, Japan, Hong Kong, and South Korea.

Americans. Researchers may use them as a first indicator as to whether response styles may contaminate the results of their cross-cultural comparisons. Moreover, although various effects display on the overall level, several areas lack general evidence and demand for further research. Table 3 tags these areas with "mixed" or "no studies". These areas display avenues for further research. Practitioners should be especially careful if they want to conduct comparisons between these groups.

3.2.1 Anglo and Asian Cultures

Studies between Anglo and Asian cultures support lower ARS (Grimm and Church 1999; Locke and Baik 2009; Watkins and Cheung 1995; Wong, Rindfleisch, and Burroughs 2003) and higher DARS (Watkins and Cheung 1995) for Anglo than Asian respondents. This difference may result form the fact that collectivistic Asian cultures support conforming behavior (Hofstede 2001) and, thus, agreeing on items regardless of the content. Contrary to these results, Riordan and Vandenberg (1994) find higher ARS for Anglos than Asians. However, their measurements focus on individual self-representation, such as a respondents' value for their employers, so that agreeing sheds a positive light on the individual. Higher ratings may, thus, not represent content-independent ARS, but SDR.

Concerning MRS and ERS, results differ for East Asians and South East Asians[7]. Chun, Campbell and Yoo (1974), Lee and Green (1991) and Zax and Takahashi (1967) find higher ERS for Anglos than East Asians. Chen, Lee, and Stevenson (1995) detect higher ERS for US Americans, but not Canadians in comparison to East Asians. Similarly, Shiomi and Loo (1999) find no differences in ERS between Japanese and Canadian students. However, they explain the lack of difference by using a student sample, which may face declining cultural differences due to their educational background. Contrary for South East Asians, Grimm and Church (1999) show lower ERS for Americans than Filipinos. Finally, Stening and Everett

---

[7] In this study South East Asia refers to Indonesia, Malaysia, Philippines, Singapore, and Thailand.

(1984) find lower ERS for Anglo than South East Asian managers, but lower ERS for Anglo compared to East Asian managers.

In line with the results for ERS, several authors find lower MRS for Anglos than East Asians (e.g., Chen, Lee, and Stevenson 1995; Lee and Green 1991; Shiomi and Loo 1999; Stening and Everett 1984; Zax and Takahashi 1967), while Stening and Everett also find higher MRS for Anglos than South East Asians. Concerning RR, Anglos show a higher tendency to circulate their answers than Asian respondents (Chun, Campbell and Yoo 1974; Watkins and Cheung 1995).

Overall, it seems that South East Asians favor extreme answers and avoid midpoints, while the opposite is true for East Asians. This difference matches the finding that Asians split up into a Far East (South East Asian countries) and a Confucian (East Asian countries) cluster (e.g., Ronen and Shenkar 2013). While acknowledgement of different believes is a key cultural trait in the Far East cluster, Confucian countries foster self-sacrificing behaviors (Gupta and Hanges 2004). Respectively, South East Asians demonstrate sincerity in their beliefs by choosing endpoints, while Confucian countries obscure their true beliefs by choosing midpoints.

Moreover, early studies of SDR show mixed evidence. Grimm and Church (1999) and Perera and Eysenck (1984) find lower SDR among Anglos than Asians. The results of Yik, Bond and Paulhus (1998) point in the opposite direction, while Heine and Lehman (1995) find no effect. A finer distinction shows that diverging effects may result from the underlying dimensions of SDR. Accordingly, Lalwani, Shavitt, and Johnson (2006) find higher scores for Anglos on ERT, but lower scores on MRT. In addition, Perera and Eysenck (1984) use the lie scale by Eysenck and Eysenck (1975), which can be interpreted as a measurement of MRT (Steenkamp, de Jong, and Baumgartner 2010). Thus, these findings may again result from the prevalence of individualism in Anglo and collectivism in Asian cultures. Individualism em-

phasizes a person's pursuit for success as basic principle, while collectivism highlights a person's contribution to a harmonic society (Hofstede 2001). For individualistic Anglo cultures, ERT, thus, reflects social desirable characteristics, while, for collectivistic Asian cultures, MRT captures social desirable characteristics. Depending on the type of SDR, both, Asians and Anglos, may show higher tendencies.

3.2.2 Other Cultural Comparisons

Few studies compare response styles between other cultures. Here, results support lower ARS, ERS and SDR for Anglos compared to Latin Americans (Clarke 2000; Ross and Mirowsky 1984). Moreover, Africans seem to show higher ARS, but lower DARS and RR than Anglos (Watkins and Cheung 1995). Mwamwenda (1993) finds higher SDR for South African women than Anglo women, while no such effect occurs for men. He measures SDR on the Eysenck Personality Inventory which covers MRT (Steenkamp, de Jong, and Baumgartner 2010). Gender differences may result from the fact that classical labels of women refer to social traits which are in line with MRT, while classical labels of men refer to ego-related traits which are in line with ERT (Hofstede 2001). In addition, Watkins and Cheung (1995) find RR to be higher for Asian than African respondents. ARS and ERS are also higher for Southern European respondents than Northwestern European respondents (Harzing 2006; Steenkamp and Baumgartner 1998; van Rosmalen, van Herk, and Groenen 2010; van Herk, Poortinga, and Verhallen 2004). Likewise, French respondents show higher ERS than Australian respondents (Clarke, 2000). Finally, Brengelmann (1959) detects that German respondents are more prone to ARS than English respondents.

3.3 Comparing Indigenes and Immigrants

3.3.1 Hispanic and Non-Hispanic White Americans

Overall, studies support higher ARS for Hispanic than non-Hispanic white Americans (Aday, Chiu, and Anderson 1980; Carr and Krause 1978; Marín, Gamba, and Marín 1992; Ross and Mirowsky 1984; Weech-Maldonado et al. 2008). Hispanics also seem to show lower DARS and higher ERS (Clarke, 2000; Hui and Triandis 1989; Johnson et al., 1997; Marín, Gamba, and Marín 1992; Weech-Maldonado et al. 2008). Nevertheless, Johnson et al. (1997) find such effect only for Mexicans, but not Puerto Ricans. In addition, Marín, Gamba, and Marín (1992) find that, for ARS and ERS, differences between groups decrease for more acculturated and educated Hispanics.

In addition, SDR seems to be more common among Hispanic than non-Hispanic White Americans (Consedine et al. 2012; Dohrenwend, 1966; Ross and Mirowsky 1984). Results by Webster (1996) show such differences only for men, but not women. Moreover, Johnson et al. (1997) again find an effect for Mexican, but not Puerto Rican Americans.

Related to these studies, Ross and Mirowsky (1984) compare Mexican Americans with Mexicans. In their study, Mexican Americans show higher SDR, but lower ARS than Mexicans in Mexico. Together with the differences in results for Mexican and Puerto Rican Americans, these results may show the influences of the acculturation process. While Puerto Rico has been under American control for a long time, no such influences on Mexico exist. Puerto Ricans may, thus, have fewer difficulties to assimilate to American culture than Mexicans. Moreover, while non-Hispanic white Americans show lower SDR than Mexicans in Mexico, Mexican Americans show even higher SDR than Mexicans in Mexico (Ross and Mirowsky 1984). This fact emphasizes that acculturation may not necessarily lead to an adaption to the host culture, but new cultural identities and behaviors may evolve.

3.3.2 African and Non-Hispanic White Americans

Studies report higher ARS, DARS, and ERS for African than non-Hispanic white Americans (Bachman and O'Malley 1984; Clarke 2000; Johnson et al. 1997). This effect may result from a higher motivation of minority respondents to show correct behavior and, thus, sincerity in their response behavior (Ross and Mirowsky, 1984) As explained, choosing the end point on both sides, is less ambiguous and, thus, supports sincerity. African Americans are, hence, stronger in agreeing, disagreeing, and extremity in their response choices. However, this effect seems to wear out with a higher education level (Dohrenwend, 1966). On the contrary, Goldsmith, White, and Stith (1987) find lower ARS among African than non-Hispanic white Americans. They use a scale by Wells (1961) with 20 self-descriptive statements. As these statements incorporate mainly negative traits, in their study, ARS might rather reflect SDR than actual ARS.

Evidence for SDR favors higher tendencies for African than non-Hispanic white Americans (Crandall, Crandall, and Katovsky 1965; Consedine et al. 2012; Goldsmith, White, and Stith 1987; Will and Verdin 1978). This effect may result from minorities' efforts to show their deference to the rules of the dominant culture (Ross and Mirowsky 1984). If minority respondents indicate a social desirable behavior, they express that they understand, apply, and appreciate the rules of the dominant culture. Contrary to these findings, Johnson et al. (1997) find no difference and the results of Dohrenwend (1966) indicate, if anything, weak support for an opposite effect. However, the latter results consider a very specific type of SDR, namely the desirability and report of medical symptoms (Carr and Krause 1978). Hence, in addition to the weak evidence, results may not be transferable to broader forms of SDR.

3.3.3 Other Comparisons between Indigenes and Immigrants

Three studies compare other groups of immigrants and indigenes. Javeline's (1999) result show higher ARS for Kazakhs in comparison to Russians in Kazakhstan. In addition, Abe and

Zane (1990) find that white Americans show higher SDR than foreign-born Asian Americans, but not to US-born Asian Americans. This result again may reflect acculturation. US-born Asians are more likely to have assimilated to white Americans than foreign-born Asians. According to Consedine et al. (2012), SDR is higher for Americans of English Caribbean, Haitian, and Dominican descent than European, Eastern European and African Americans. Moreover, SDR is higher for African than Eastern European Americans and Dominican than Haitian immigrant women.

3.4 Comparisons along Cultural Dimensions

The review of the third research stream splits along the dimensions individualism, masculinity, power distance, and uncertainty avoidance as defined in table 2 above. For each dimension, the review examines studies that work with the cultural dimension and then compares results to the results of cross-country comparisons and comparisons of immigrants and indigenes.

Table 4 gives an overview of the results for these four dimensions. In several areas, scarce and mixed results clearly demand for further research. Table 4 tags these areas with the labels "mixed" and "no studies". Again, practitioners should treat comparisons among these dimensions with care, while researchers should take them as avenues for further research.

Table 4: Effects of Cultural Dimensions on Selected Response Style[1]

	Individualism	Masculinity	High Power Distance	High Uncertainty Avoidance
Acquiescence Response Style	(-)	(-)	*mixed*	*mixed*
Disacquiescence Response Style	(+)	*no studies*	(-)	*no effect*
Extreme Response Style	*mixed*	*mixed*	(+)	(+)
Midpoint Response Style	(-)	*no studies*	(-)	*no effect*
Response Range	*no effect*	*no effect*	*no effect*	*no effect*
Egoistic Response Tendency	*mixed*[2]	(+)	*no studies*	*no studies*
Moralistic Response Tendency	(-)[2]	(-)	(+)	*no effect*

1) cells include only response styles that researchers have analyzed; 2) in a study by Lalwani et al. (2006), effects only display for horizontal, but not vertical forms of collectivism and individualism.

3.4.1 Individualism and Collectivism

For the sake of social harmony, individuals in collectivistic cultures are more likely to conform (Hofstede 2001) and, hence, more likely to agree on an item. Accordingly, empirical studies show lower ARS among individualistic than collectivistic cultures, while the opposite holds for DARS (Davis, Resnicow, and Couper 2011; Johnson et al. 2005; Harzing 2006[8]; Hofstede 2001). Results fit to cross-country comparison, where ARS is lower and DARS is higher for individualistic Anglo than collectivistic Asian countries and individualistic North-Western European than collectivistic Southern European states.

For ERS, results are mixed. Different authors find individualism to induce more (de Jong et al. 2008; Chen, Lee, and Stevenson 1995), similar (Johnson et al., 2005), and less (Davis, Resnicow, and Couper 2011; Peterson, Rhi-Perez, and Albaum 2014) ERS than collectivism.

---

[8] Harzing (2006) defines effects for negative and positive extreme response styles. As these two measures do not consider extreme response style in general, they rather represent acquiescence and disacquiescence bias.

These mixed results gather support from cross-country studies. These studies find more extreme answers for Southern Europeans, Latin Americans and South East Asians (collectivistic) in comparison to North-Western Europeans and Anglos (individualistic), while this pattern is reversed for collectivistic East Asians.

Chen, Lee, and Stevenson (1995) and Harzing (2006) find that MRS decreases with stronger endorsement of individualism. Supportingly, cross-country comparisons show higher MRS for individualistic Anglos than collectivistic East Asians. Here, although midpoints are more ambiguous than extreme answers, they may also avoid standing out from others and, hence, help to deemphasize individualistic tendencies. However, results also show that the individualism-collectivism distinction is not sufficient to explain all cross-country comparisons. Namely, comparisons of individualistic Anglo and collectivistic South East Asian cultures show an opposite trend to the overall effect of individualism on MRS. Moreover, only Peterson, Rhi-Perez, and Albaum (2014) assess the RR for individualistic versus collectivistic cultures. They find no correlation between individualism and RR.

As for SDR, effects are more diversified. On a general level, Steenkamp, de Jong, and Baumgartner (2010) find higher ERT and MRT for collectivistic cultures and results by van Hemert et al. (2002) and Smith (2004) support higher MRT for collectivistic cultures. On a more detailed level, the distinction of horizontal and vertical forms of individualism and collectivism may add explanation. While horizontal individualism leads to higher ERT, horizontal collectivism encourages MRT (Lalwani, Shavitt, and Johnson 2006). No tendencies evolve for vertical forms of individualism and collectivism. Result fit to the evidence and explanation for cross-country comparisons between collectivistic Asian and individualistic Anglos cultures. Overall, both individualism and collectivism bear the potential for SDR (Steenkamp, de Jong, and Baumgartner 2010). Individualism encourages ERT, while collectivism promotes MRT.

3.4.2 Masculinity and Femininity

As femininity connects to submissive behavior, ARS is less common in masculine than feminine cultures (Davis, Resnicow, and Couper 2011; Johnson et al. 2005). Evidence for ERS is mixed. Johnson et al. (2005) and de Jong et al. (2008) find higher ERS for masculine cultures, Peterson, Rhi-Perez, and Albaum (2014) no correlation of ERS with masculinity, and Davis, Resnicow, and Couper (2011) marginal support for lower ERS in masculine cultures. Moreover, Peterson, Rhi-Perez, and Albaum (2014) find no correlation of masculinity with the RR.

Concerning SDR, masculine cultures seem to support ERT, while feminine cultures seem to support MRT (Steenkamp, de Jong, and Baumgartner 2010). Masculine cultures emphasize a person's power and success (de Jong et al. 2008). Individuals, hence, seek for possibilities to enhance their own image, which leads to ERT. On the contrary, female cultures stress caring, good relationships, and quality of life (Hofstede 2001). Respondents in these cultures strengthen MRT. This result finds some support in the cross-country comparisons above. Perera and Eysenck (1984) show higher MRT for Sri Lanka as a female culture than the UK as a masculine culture. Nevertheless, van Hemert et al. (2002) and Smith (2004) find no significant effect for masculinity on SDR in their studies.

3.4.3 Power Distance

Concerning power distance, Hofstede (2001) and Harzing (2006) find a positive effect on ARS and weak support for a negative effect on DARS. These effects may reflect the high importance of rules and authorities in high power distance cultures (Hofstede 2001). They may force individuals to conform and, thus, agree on items. On the contrary, Johnson et al. (2012) find a negative effect on ARS. They explain this result by means of measurement. While they measure power distance at the individual respondent level, other studies have used

national-level data. If they repeat their analysis on the national level, results also support a positive effect of power distance on ARS.

Moreover, although de Jong et al. (2008) find no effect, a high power distance seems to promote ERS (Johnson et al. 2005; Peterson et al., 2014). Strong authorities stress decisiveness in cultures with a high power distance (Johnson et al. 2005). As extreme response categories state clear answers and avoid ambiguity (Baumgartner and Steenkamp 2001), power distance enforces ERS. This result gets support from cross-country comparisons. High power distant Asian and Latin American cultures show higher ERS than low power distant Anglo cultures. In line with these effects, Harzing (2006) finds a negative effect of high power distance on MRS. In addition, no significant correlation evolves between power distance and the RR (Peterson, Rhi-Perez, and Albaum 2014).

Finally, a higher power distance seems to promote MRT (Smith 2004; van Hemert et al. 2002). Individuals in high power distant cultures accept unequal distributions of power and express this acceptance in their actions (Hofstede 2001). Thus, they should point out how they conform to the rules of society, which is in line with the definition of MRT. In support of this result, Asians (higher power distance) show higher MRT than Anglos (low power distance) in cross-country comparisons.

3.4.4 Uncertainty Avoidance

Finally for uncertainty avoidance, evidence is mostly mixed or shows no effect. So Johnson et al. (2005) reveal a negative connection to ARS, while Harzing (2006) finds an opposite effect. Moreover, Harzing (2006) shows no effect of uncertainty avoidance on DARS and MRS. Peterson et al. (2014) find no correlation of uncertainty avoidance and RR. Finally, uncertainty avoidance shows no connection to MRT (Smith, 2004; van Hemert et al. 2002). Only for ERS, results seem to favor a positive effect of uncertainty avoidance (de Jong et al. 2008; Peterson, Rhi-Perez, and Albaum 2014). This effect may result from the higher certain-

ty that extreme answers express compared to ambiguous midpoints (Baumgartner and Steenkamp 2001). Nevertheless, Johnson et al. (2005) find also no effect of uncertainty avoidance on ERS.

3.5 Studies considering the Moderation of Item Characteristics on the Effect of Culture on
 Response Style

As explained above, items obtain specific content and measurement characteristics. While content-related factors may only affect content-dependent response styles, measurement format may impact other response styles across cultures. Several authors have found that item characteristics may moderate the effect of culture on response styles. Thus for each characteristic, the review below first provides a brief conceptual consideration and then assesses empirical results. As a result, table 5 sums up key findings for reach characteristic.

Table 5: Key Findings for Influences of Item Design on the Effect of Culture on Response Styles

Item Characteristic	Key Findings
Content-related characteristics	• cultures differ in their tendencies to apply egoistic response tendencies and moralistic response tendencies for different topics
Language	• extreme response styles are lower in a first compare to a second language • response ranges are narrower in a first compared to a second language • mid response styles are higher in a first compared to a second language
Item polarity	• combining reversed and non-reversed items for a scale may reduce acquiescence response styles for Anglo, but not for Asian cultures • lower acquiescence response styles for Anglo compared to Asian respondents persist for scales that consist of items with similar polarity
Scale format	• interrogative scales can reduce differences in acquiescence response style between Anglo and Asian cultures
Number of answer categories	• in most of the cases, lower numbers of answer categories reduce differences in extreme response styles • lower numbers of answer categories seem to increase differences in acquiescence response styles
Labeling of answer categories	• sparse evidence for no effect of labeling all answer categories in comparison to label only endpoint categories on differences in response styles between cultures • intensity and familiarity of answer labels differ between cultures • higher familiarity of extreme answer categories in a culture may lead to higher extreme response styles

3.5.1 Content-related Characteristics

Different item contents may lead to different perception of socially desirable answers across respondents. Hereby, definitions of social desirability rely on norms that may vary

across sociocultural groups (de Jong et al. 2010). In this manner, studies show that content-related characteristics influence differences in SDR across cultures.

Dohrenwend (1966) finds African, Irish and Jewish Americans to evaluate several symptoms of psychological disorder as less desirable than Puerto Ricans. Likewise in a study by Yik, Bond, and Paulhus (1998), Chinese students show generally lower self-enhancement in comparison to North American students, but differences in self-enhancement between both samples vary strongly with the underlying scale. In addition, Steenkamp, de Jong, and Baumgartner (2010) analyze established scales on their proneness to SDR across countries. They find that scales differ in their contamination through ERT and MRT within and between countries. For example, health consciousness shows contamination by MRT, but not ERT in the USA, while, in Italy, only ERT, but not MRT does contaminate values for this scale. Overall, results support that cultures do not only differ in their general tendencies to apply ERT or MRT, but these differences vary across item content.

3.5.2 Language

The role of language is of particular interest, when researchers study bi- or multilinguals. Here, it is the case that different languages may activate different cultural frameworks (e.g., Lau-Gesk 2003). Respectively, interpretations and answers may generally differ between first and second languages (e.g., Luna, Ringberg, and Perracchio 2008). In detail, anchor contraction effects may cause that reference points are more intense in a first than a second language. As more intense answer categories represent more extreme opinions, respondents are less likely to choose these answer categories (Weijters, Geuens, and Baumgartner 2013).

Studies support that respondents experience labels in their first language as more intense than in a second language. Harzing (2006) and McCrae (2002) find higher ERS, lower MRS, and lower RR in a second compared to a first language. Additional results by de Langhe et al. (2011) show higher ERS for emotional items in a second language. Due to their higher inten-

sity, respondents actually avoid extreme answers in their first language. They stick to answer categories close to the midpoint. Accordingly, ERS and RR are lower and MRS is higher.

3.5.3 Item Polarity

Throughout a survey, researchers may use reversed and non-reversed items. While non-reversed items are in line with the scales initial meaning, reversed items include questions or statements that are opposite to this meaning (Weijters and Baumgartner 2012). For Western cultures, using both types of items forces respondents to choose opposite answers because they do not tolerate contradictions (Peng and Nisbett 1999). In contrast, other cultures, such as Buddhist cultures, openly accept or even promote agreeing with contradictory statements.

Here, Wong, Rindfleisch, and Burroughs (2003) show that, while using reversed and non-reversed items reduces ARS for American respondents, this effect is weaker or even non-existent among Asian respondents. In addition, Grimm and Church (1999) provide evidence that higher ARS for Asians also exists when items reveal a lower level of opposing polarity. They find only a low level of negative correlations across items in a measurement of individualism and collectivism by Hui (1988). Nevertheless, results still indicate higher ARS for Filipino respondents in comparison to American respondents for this scale. Both studies thus provide initial evidence that item polarity may influence differences in response styles across cultures. Yet no study assess whether results actually differ in the percentage of reversed items for a scale.

3.5.4 Scale Format

The scale format influences how respondents interpret and rate items. For example, a standard Likert format asks for the respondents' agreement with an item, while an interrogative format asks for a decision between two statements. Accordingly, respondents are able to agree with two contradicting statements on a Likert scale, but cannot do so, when researchers

directly oppose these statements in one interrogative item (Wong et al., 2003). As tendencies to agree with contradicting statements vary across cultures, differences in response styles between cultures should also vary across these scale formats. For other scale formats, similar differences are plausible.

For seven scale formats, Crosby (1969) tests whether French-Canadians differ from English-Canadians in their response styles. Across all formats, he finds no differences in ARS between both groups. Contrarily, Wong, Rindfleisch, and Burroughs (2003) identify significant differences for American and Asian respondents. When using Likert scales Asian respondents show higher agreeability with contradictive items than Americans, while this difference does not exist for interrogative scales. In a study by Javeline (1999), a similar scale type is able to reduce differences in ARS between Kazakh and Russian respondents. For the latter studies, the difference to the findings of Crosby (1969) may result from the fact that cultural differences in Crosby's study are much lower than in the other studies. In fact, Crosby uses two populations that share a common cultural heritage. Such data may reflect language, but no cultural differences.

3.5.5 Number of Answer Categories

When individuals respond to a survey item, they have to choose among answer categories provided by the researcher. At the same time, their response may not fit to this scheme, but may exist in subjective categories of judgment (Hui and Triandis 1989). Thus, respondents need to match these subjective categories to the answer categories provided. Two issues are prevalent in this context.

On the one hand, respondents may have to match several, detailed subjective categories to one answer category in the questionnaire. Here, certain answer categories provided may cover more subjective categories of judgment than other answer categories provided. As respondents choose these answer categories more often, response styles arise. In this context,

176

cultures differ in the level of detail that subjective judgments reflect (D'Andrade 1972; Wright et al. 1978) and, thus, the prevalence of matching problems. Here, a larger number of answer categories provided may, generally, reduce matching problems and corresponding response styles (Hui and Triandis 1989). More answer categories are likely to even capture more detailed subjective judgment schemes, while, for less detailed subjective judgment schemes, no difference exists for more or less answer categories provided.

On the other hand, respondents have to interpret the given answer scheme. While this interpretation is straightforward for a low number of answer categories, more answer categories may promote various subjective interpretations across respondents (Cronbach 1950). This subjectivity may introduce systematic variation and, thus, response styles in respondents' answers. Here, interpretations of answer schemes may vary with cultural backgrounds (e.g., de Langhe et al. 2011).

Similarly to the divergence of both issues, no clear picture evolves for studies that analyze how the number of answer categories moderates cultural influences on response styles. For ERS, overall evidence seems to favor less over more answer categories. Clarke (2000), Grimm and Church (1999), and Marín, Gamba, and Marín (1992), find lower differences in ERS for low compared to high numbers of answer categories. Their samples allow them to conduct several cultural comparisons across and within countries and, thus, support generalizability. On the contrary, the results of Hui and Triandis (1989) show an opposite effect. Likewise, higher numbers of answer categories also seem to decrease differences in ARS (Grimm and Chruch 1999). So no clear recommendation depicts on how researchers can chose the number of answer categories to reduce differences in response styles.

3.5.6 Labeling of Answer Categories

Greenleaf (1992) indicates that labeling all answer categories may cause different levels of response styles in comparison to labeling only endpoints. Labeling a response category

reduces its ambiguity (French-Lazovik and Gibson 1984; Weijters, Cabooter, and Schillewaert 2010). Therefore, when labeling only end categories, they appeal less ambiguous than other categories, while, when labeling all categories, no difference in ambiguity between answer categories should exist. In this context, choosing ambiguous categories contradicts respondents' pursuit to express sincerity. Thus, if they aim to show sincerity in their answers, they may focus on unambiguous answer categories. They choose these categories more often, which leads to respective response styles. Culture may moderate this effect because certain cultures especially value sincerity in respondents' answers, while this value is less pronounced in other cultures (e.g., Hui and Triandis 1989).

De Langhe et al. (2011) compare the impact of labeling all answer categories and labeling only endpoints on response styles in a first and second language. They find no difference for respondents' first in comparison to their second language. In line with this study, Weijters, Cabooter, and Schillewaert (2010) assess scale labeling in two separate studies, one study with Dutch-speaking and one study with English-speaking respondents. Although they do not directly test for cultural differences, they infer no differences because results from the Dutch-speaking respondents are predictive for the results of English-speaking respondents. Nonetheless, no study so far directly compares full and endpoint labeling across cultures.

In addition, respondents are less likely to choose answer categories with a more intense label because they perceive these answer categories as more extreme (de Langhe et al. 2010). At the same time, they are more likely to choose familiar answer categories because they reduce the ease of processing (Weijters, Geuens, and Baumgartner 2013). In this context, intensity of labels differ across cultures. Voss et al. (1996) show that Japanese equivalents for the English labels "good" and "very good" are more intense. They translate English labels via means of translation-back-translation. While English-speaking respondents indicate their relative magnitude in comparison to a neutral category with 74 and 87, Japanese respondents rate

them 91 and 101. These variations in the intensity of labels may, thus, lead to differences in response styles across cultures. Similarly, the choice of more familiar labels is a systematic deviation from true scores and, thus, a response style. Like intensity, familiarity of labels differs across cultural groups. For example, Weijters, Geuens, and Baumgartner (2013) find the label "Entièrement d'accord" to be more familiar among French-speaking respondents in France in comparison to French-speaking respondents in Canada. They also show that differences in ERS result from differences in the familiarity of endpoint labels across cultures.

4 Implications for Researchers and Research Gaps

This paper derives generalizations on influences of culture on response styles. These generalizations demonstrate that culture is a valid issue that may harm questionnaire results and corresponding conclusions. Researchers, thus, need to assess their sample composition for potential culturally induced response styles. They can use generalizations from this study to control for potential contamination of survey measurements. Tables 3 and 4 provide particular insights that help identify potential issues in the sample structure. For example, if researchers find a result across a cultural diverse sample, they can use these tables to check whether this result may be an artifact from differences in response styles across cultures in the sample.

This study also depicts areas where evidence is sparse or mixed. Here, researchers should establish further research. In particular, the seven issues below demand for further studies and offer promising avenues for further research.

4.1 Research Gap 1: Culture and Scale Format

Only Crosby (1969), Javeline (1999), and Wong, Rindfleisch, and Burroughs (2003) have assessed response styles across cultures for different scale formats. The latter two studies assess only Likert and interrogative studies and provide evidence for an influence of the scale format of cultural variation in response styles. Crosby (1969) assesses several bi- and unipolar

scales, such as a pictorial facial scale, a semantic differential, and a pictorial hands scale. He finds no effect of response styles across all scale types. However, his sample reveals only marginal differences in cultural backgrounds. Thus, these scale formats may as well reveal an effect on the variation of response styles across more diverse cultural groups. Especially, comparisons among (1) bi- and unipolar scales and (2) pictorial and non-pictorial scales are promising avenues for further research. For bi- and unipolar scales, agreement intensity may differ between both formats and pictorial elements may be prone to different interpretations across cultures.

4.2 Research Gap 2: Interplay of Acculturation and Scale Formats

A second research gap exists on how differences in response styles due to acculturation vary with the underlying scale format. On the one hand, acculturation studies only address differences in response styles based on home and host culture, but do not compare this effect across different scale formats (e.g., Davis, Resnicow, and Couper 2011). On the other hand, when researchers compare response styles for different scale formats, they only do so across cultures, but leave out the acculturation perspectives (e.g., Wong, Rindfleisch, and Burroughs 2003). However, immigrants do not only incorporate cultural elements of the home and host culture, but may evolve patterns that differ from both (e.g., Wallendorf and Reily 1983). Thus, research needs to assess differences as to how acculturation patterns that are independent to underlying cultures shape response styles dependent on the scale format.

4.3 Research Gap 3: Labeling All Answer Categories versus Labeling Only Endpoint Categories

As explained above, theoretical discourse suggests an influence of labeling all answer categories in comparison to labeling only endpoint categories on the effect of culture for response styles. Moreover, Weijters, Cabooter, and Schillewaert (2010) explicitly point to such

180

effect. Nonetheless, no study so far analyzes whether cultural variance in response styles exists between both cases. Furthermore, it would be especially important to check whether the numerous recommendations to use endpoint or full labeling as by Weijters, Cabooter, and Schillewaert (2010) still hold with more culturally diverse samples.

4.4 Research Gap 4: Global Culture and Differences in response Styles

No study so far assesses how global cultural trends, as, for example, shown by Kjelgaard and Askegaard (2005), and their prevalence influence the effect of culture on response styles. At the same point, globalization increases the impact of global cultural trends on the individual. Individuals are, thus, more likely to take over elements from global culture. As these elements are the same across traditional cultures, differences in response styles across formerly distinct cultural groups may decrease. This process may, for example, explain the results of Shiomi and Loo (1999), which show no differences on ERS for high-educated Canadian and Japanese students. As students associate with a highly international environment, they may broadly get in touch with global culture. Meanwhile, respondents across and within countries differ in their degree of receptivity of global culture. For example, individuals in emerging markets may be less receptive than individuals in industrialized markets (Kumar et al. 2013). Here, global culture may also lead to higher differences between and within cultures. Future research should, hence, analyze whether and how global culture may change cultural patterns for response styles.

4.5 Research Gap 5: Additional Distinctions of Cultural Dimensions

While some studies exist that consider the effect of generic cultural dimensions on response styles, evidence is often mixed or even missing. Here, a finer distinction of cultural dimensions may help solve some of the contradictions of effects across studies. Accordingly, Lalwani, Shavitt, and Johnson (2006) show that the distinction between horizontal and verti-

cal forms of individualism and collectivism may add to the explanation of differences in effects on ERT and MRT. Davis, Resnicow, and Couper (2011) demonstrate that a finer distinction of masculinity and femininity may provide further insights about cultural effects on ARS and ERS. Johnson et al. (2012) indicate that different results may occur, if researchers measure cultural dimensions at the individual instead of the national level. In addition, other conceptualizations of cultural values, beliefs, and norms may add explanation for cultural differences in response styles. Here, Steenkamp, de Jong, and Baumgartner (2010) find effects for the cultural values by Schwartz (1992) on ERT and MRT. Authors should, hence, analyze how a (1) finer distinction of cultural dimensions, (2) an individual-based measurement of these dimensions, (3) the use of other generic cultural dimensions, such as those by Schwartz (1992), and (4) the interplay of these dimensions with individual factors may explain response styles.

4.6 Research Gap 6: Business Culture and Response Styles

The definition of culture is not limited to an ethnic or national perspective, but refers to other sociocultural classifications like business culture. Similar to an ethnic definition of culture, business culture splits among several cultural dimensions like in-group collectivism or assertiveness (Hanges and Dickson 2004). In this context, Harzing (2006) actually shows influences of these dimensions on some response styles. She uses data on a country-level. However, dimensions of business culture may also vary between markets and companies within a country. As questionnaires are a common tool to study business phenomena (e.g., Brown et al. 2011), future studies should analyze the impact of business culture on the response styles of individual managers in detail. They may even combine the traditional and business culture perspective and assess how both types of cultures interact in their influence on response styles.

4.7 Research Gap 7: Replicability of Effects with Marketing Constructs

Although several studies exist that look for cultural contamination on marketing constructs, the main part of study comes from psychological or health care research. These studies partly involve sensitive issues, such as health questions, for which consumers reveal differences in attitudes across cultures (Johnson et al., 1997). Moreover, Steenkamp, de Jong, and Baumgartner (2010) show substantial variance of at least ERT and MRT across cultures and contents. Thus, further studies have to assess (1) whether results are replicable for marketing constructs and (2) how cultural effects on response styles differ across other established marketing scales.

5   Implications for Managers

If managers conduct surveys, they have to consider that underlying cultures shape response styles and, thus, results. This fact is especially critical, if managers intend to conduct cross-cultural comparisons. However, cultural influences do not only affect cross-cultural studies, but lead to contamination in other studies. In particular, studies between immigrants and indigenes show that response styles differ between cultural groups within a country. Thus, it is inevitable for managers to assess the cultural backgrounds of respondents even in studies that do not refer to cross-cultural issues. Otherwise, their results are at the risk to deflate, inflate, or even converse effects.

Table 3 and 4 show some generalizations from findings for response styles across cultural and migrant groups. Managers can use these generalizations as indicators for potential contamination in their findings. Therefore, first, they need to collect data on the cultural backgrounds of their respondents. Second, if they know these backgrounds, they can, check whether these fit to groups in table 3 and derive corresponding differences in response styles. Third, if no such comparison exists, they can take table 4 in consideration. To use this table,

they have to work out which cultural dimensions show differences between groups in their sample. Here, they can use index scores for the dimensions as provided by Hofstede (2016). For dimensions that show differences, they can check table 4 for generalizable differences in response styles. Fourth, they can compare both, effects on response styles from table 3 an 4, to the effects they find in their study. The following suggestions may help identify issues: If differences in response styles are in line with an effect, effects may be an artifact of response styles and actually weaker or non-existent. If differences in response styles are opposite to an effect, actual effects may be even stronger. If no effect exists, response styles may counteract an actual effect.

In addition, for SDR, standard tests show mixed and partly contradictory results. A solution to this issue seems to lie in the separation of ERT and MRT. As shown above, results are much clearer for this distinction. So, if managers conduct sensitive questionnaires with different cultural groups, they should check for both kinds of SDR separately.

As the conceptual framework of this study shows, item design is crucial when researchers expect culturally induced response styles. Here, studies support several moderating effects. Managers, hence, need to consider potential sample compositions when they construct items. For this purpose, table 5 sums up several results that provide guidance for item design. For example, if researchers intend to study respondents from Asia, interrogative scale formats may be more advantageous than standard Likert scales. Moreover, although studies provide no clear recommendation on the optimal number of answer categories, both too much and too few answer categories may enhance certain response styles. Thus, standard choices, such as 5 or 7 options, seem most reasonable.

6   Conclusion

In a broad literature review, this paper assesses how culture shapes response styles. For this purpose it first provides a general framework that explains how response styles arise and

how culture interferes with this process. Culture in this context is not restricted to a person's initial beliefs, values, and norms, but varies with external influences, such as global cultural trends. Based on this framework, the paper compares empirical studies that analyze how culture influences response styles. Overall, studies support that culture shapes response styles and that cultural differences for response styles may contaminate results of statistical analysis. Importantly, this contamination is not restricted to comparisons between countries, but also appears for comparisons of immigrants and indigenes. Thus, results may experience contamination even in studies that researchers do not conduct at a cross-cultural level. Taken together, studies show some generalizable effects across cultural groups. For instance, a broad consensus exists that individualistic and collectivistic cultures promote different kinds of SDR. In addition, the study also reviews potential moderator effects through item design. Here some generalizable evidence exists. For example interrogative scales may reduce issues with ARS. Nevertheless, many topics remain unsolved or show mixed evidence. For instance, no clear recommendation exists on the number of answer categories to use. Therefore, this review derives several research gaps that provide promising avenues for further research. The review closes by transferring results into actionable recommendations for practice and derives management implications.

REFERENCES PROJECT 4

Abe, Jennifer S. and Nolan W. S. Zane (1990). "Psychological Maladjustment among Asian and White American College Students: Controlling for Confounds," *Journal of Counseling Psychology*, 37 (4), 437-444.

Aday, Lu A., Grace Y. Chiu, and Ronald Andersen (1980). "Methodological Issues in Health care Surveys of the Spanish Heritage Population," *American Journal of Public Health*, 70 (4), 367-374.

Askegaard, Søren, Eric J. Arnould, and Dannie Kjeldgaard (2005). "Postassimilationist Ethnic Consumer Research: Qualifications and Extensions," *Journal of Consumer Research*, 32 (1), 160-170.

Bachman, Jerald G. and Patrick M. O'Malley (1984). "Yea-saying, Nay-saying, and Going to the Extremes: Black-white Differences in Response Styles," *Public Opinion Quarterly*, 48 (2), 491-509.

Baumgartner, Hans and Jan-Benedict E. M. Steenkamp (2001). "Response Styles in Marketing Research: A Cross-national Investigation," *Journal of Marketing Research*, 38 (2), 143-156.

Berry, John W. (1986). "Multiculturalism and Psychology in Plural Societies," in L. H. Ekstrand (Ed.), *Ethnic Minorities and Immigrants in a Cross-cultural Perspective*, Albasserdam: Swets and Zeitlinger, 35-51.

Brakus, J. Joško, Bernd H. Schmitt, and Lia Zarantello (2009). "Brand Experience: What is It? How is It Measured? Does it Affect Loyalty?," *Journal of Marketing*, 73 (3), 52-68.

Brengelmann, Johannes C. (1959). "Differences in Questionnaire Responses between English and German Nationals," *Acta Psychologica*, 16, 339-355.

Broen, William E. and Robert D. Wirt (1958). "Varieties in Response Sets," *Journal of Consulting Psychology*," 22 (3), 237-240.

Brown, Brian P., Alex R. Zablah, Danny N. Bellenger, and Wesley J. Johnston (2011). "When do B2B Brands Influence the Decision Making of Organizational Buyers? An Examination of the Relationship between Purchase Risk and Brand Sensitivity," *International Journal of Research in Marketing*, 28 (3), 194-204.

Buzova, Daniela, Silvia Sanz-Blas, and Amparo Cervera-Taulet (2016). "Cross-cultural Perceptions of Onshare Guided Tours: A Qualitative Approach Based on eWOM," *Psychology and Marketing*, 33 (12), 1054-1061.

Carr, Leslie G. and Neal Krause (1978). "Social Status, Psychiatric Symptomatology, and Response Bias," *Journal of Health and Social Behavior*, 19 (1), 86-91.

Chattaraman, Veena, Sharron J. Lennon, and Nancy A. Rudd (2010). "Social Identity Salience: Effects on Identity-based Brand Choices of Hispanic Consumers," *Psychology and Marketing*, 27 (3), 263-284.

Chen, Chuansheng, Shin-Ying Lee, and Harold W. Stevenson (1995). "Response Style and Cross-cultural Comparisons of Rating Scales among East Asian and North American Students," *Psychological Science*, 6 (3), 171-175.

Chun, Ki-Taek, John B. Campbell, and Jong H. Yoo (1974). "Extreme Response Style in Cross-cultural Research. A Reminder," *Journal of Cross-Cultural Psychology*, 5 (4), 465-480.

Church, Timothy A. (2000). "Culture and Personality: Toward an Integrated Cultural Trait Psychology," *Journal of Personality*, 68 (4), 651-703.

Clarke, Irvine III (2000). "Extreme Response Style in Cross-cultural Research: An Empirical Investigation," *Journal of Social Behavior and Personality*, 15 (1), 137-152.

Consedine, Nathan S., Carol Magai, David Horton, and William M. Brown (2012). "The Affective Paradox: An Emotion Regulatory Account of Ethnic Differences in Self-reported Anger," *Journal of Cross-Cultural Psychology*, 43 (5), 723-741.

Crandall, Virginia C., Vaughan J. Crandall, and Walter Katovsky (1965). "A Children's Social Desirability Questionnaire," *Journal of Consulting Psychology*, 29 (1), 27-36.

Cronbach, Lee J. (1950). "Further Evidence on Response Sets and Test Design," *Educational and Psychological Measurement*, 10 (1), 3-31.

Crosby, Richard W. (1969). "Attitude Measurement in a Bilingual Culture," *Journal of Marketing Research*, 6 (4), 421-426.

D'Andrade, Roy G. (1972). "Categories of Disease in American-English and Mexican-Spanish," in R. N. Shepard, A. K. Romney, S. B. Nerlove (Eds.), *Multidimensional Scaling*, New York: Seminar Press, 11-54.

D'Andrade, Roy G. (1995). *The Development of Cognitive Antropology*, Cambridge: University Press.

Davis, Rachel E., Ken Resnicow, and Mick P. Couper (2011). "Survey Response, Acculturation, and Culture among a Sample of American Adults," *Journal of Cross-Cultural Psychology*, 42 (7), 1219-1236.

De Jong, Martijn G., Jan-Benedict E. M. Steenkamp, Jean-Paul Fox, and Hans Baumgartner (2008). "Using Item Response Theory to Measure Extreme Response Style in Marketing: A Global Investigation," *Journal of Marketing Research*, 45 (1), 104-115.

De Jong, Martijn G., Rik Pieters, and Jean-Paul Fox (2010). "Reducing Social Desirability Bias through Item Randomized Response: An Application to Measure Underreported Desires," *Journal of Marketing*, 47 (1), 14-27.

De Langhe, Bart, Stefano Puntoni, Daniel Fernandes, and Stijn M. J. van Ossealaer (2011). "The Anchor Contraction Effect in International Marketing Research," *Journal of Marketing Research*, 48 (2), 366-380.

Dohrenwend, Bruce P. (1966). "Social Status and Psychological Disorder: An Issue of Substance and an Issue of Method," *American Sociological Review*, 31 (1), 14-34.

Eysenck, Hans J. and Sybil B. G. Eysenck (1975). *Manual of the Eysenck Personality Questionnaire*, San Diego: Educational and Industrial Testing Service.

Faber, Ronald J., Thomas C. O'Guinn, and John A. McCarty (1987). "Ethnicity, Acculturation, and the Importance of Product Attributes," *Psychology and Marketing*, 4 (2), 121-134.

Fischer, Marc, Franziska Völckner, and Henrik Sattler (2010). "How Important are Brands? A Cross-category, Cross-country Study," *Journal of Marketing Research*, 47 (5), 823-839.

French-Lazovik, Grace and Curtis L. Gibson (1984). "Effects of Verbally Labeled Anchor Points on the Distributional Parameters of Rating Measures," *Applied Psychological Measurement*, 8 (1), 49-57.

Goldsmith, Ronald E., J. Dennis White, and Melvin T. Stith (1987). "Values of Middle-class Blacks and White: A Replication and Extension," *Psychology and Marketing*, 4 (2), 135-144.

Green, Kathy E. (1996). "Sociodemographic Factors and Mail Survey Response," *Psychology and Marketing*, 13 (2), 171-184.

Greenleaf, Eric A. (1992). "Measuring Extreme Response Bias," *Public Opinion Quarterly*, 56 (3), 328-351.

Grimm, Stephanie D. and Timothy A. Church (1999). "A Cross-cultural Study of Response Biases in Personality Measures," *Journal of Research in Personality*, 33 (4), 415-441.

Gupta, Vipin and Paul J. Hanges (2004). "Regional and Climate Clustering of Societal Cultures," in R. J. House, P. J. Hanges, M. Jarvidan, P. W. Dorfmann, V. Gupta (Eds.), *Culture, leadership, and organization*, Thousand Oaks et al.: Sage, 178-218.

Hanges, Paul J. and Marcus W. Dickson (2004). "The Development and Validation of the GLOBE Culture and Leadership Scales," in R. J. House, P. J. Hanges, M. Jarvidan, P. W. Dorfmann, V. Gupta (Eds.), *Culture, leadership, and organization*, Thousand Oaks et al.: Sage, 122-151.

Harzing, Anne-Wil (2006). "Response Styles in Cross-national Survey Research. A 26-Country Study," *International Journal of Cross Cultural Management*, 6 (2), 243-266.

Heine, Steven J. and Darrin R. Lehman(1995). "Social desirability among Canadian and Japanese students," *Journal of Social Psychology*, 135 (6), 777-779.

Hofstede, Geert (2001). *Culture's Consequences*, Thousand Oaks et al.: Sage.

Hofstede, Geert (2016). *Cultural Dimensions* [available at http://geert-hofstede.com/countries.html, accessed November 13, 2016].

Hui, C. Harry (1988). "Measurement of Individualism-collectivism," *Journal of Research in Personality*, 22 (1), 17-36.

Hui, C. Harry, Triandis, Harry C. (1989). "Effects of Culture and Response Format on Extreme Response Style," *Journal of Cross-Cultural Psychology*, 20 (3), 296-309.

Javeline, Debra (1999). "Response Effects in Polite Cultures. A Test of Acquiescence in Kazakhstan," *Public Opinion Quarterly*, 63 (1), 1-28.

Ji, Li-Jun, Norbert Schwarz and Richard E. Nisbett (2000). "Culture, Autobiographical Memory, and Behavioral Frequency Reports: Measurement Issues in Cross-cultural Studies," *Personality and Psychological Bulletin*, 26 (5), 585-593.

Johnson, Timothy, Patrick Kulsea, Young I. Cho, and Sharon Shavitt (2005). "The Relation between Culture and Response Styles. Evidence from 19 Countries," *Journal of Cross-Cultural Society*, 36 (2), 264-277.

Johnson, Timothy, Diane O'Rourke, Noel Chavez, Seymour Sudman, Richard Warnecke, Loretta Lacey, and John Horm (1997). "Social Cognition and Responses to Survey Questions among Culturally Diverse Populations," in L. Lyberg, P. Biemer, M. Collins, E. de Leeuuw, C. Nippo, N. Schwarz, D. Trewin (Eds.), *Survey Measurement and Process Quality*, New York: Wiley, 87-113.

Kim, Claire H. and Youjae Yi (2016). "The Effects of Impression Management on Coupon Redemption across Cultures," *Psychology and Marketing*, 33 (7), 573-583.

Kjeldgaard, Dannie and Søren Askegaard (2006). "The Glocalization of Youth Culture: The global Youth Segment as Structures of Common Difference," *Journal of Consumer Research*, 33 (2), 231-247.

Kumar, V., Amlesh Sharma, Riddhi Shah, and Bharath Rajan (2013). "Establishing Profitable Customer Loyalty for Multichannel Companies in the Emerging Economies: A Conceptual Framework," *Journal of International Marketing*, 21 (1), 57-80.

Lalwani, Ashok K., Sharon Shavitt, and Timothy Johnson (2006). "What is the Relation between Cultural Orientation and Socially Desirable Responding?," *Journal of Personality and Social Psychology*, 90 (1), 165-178.

Lau-Gesk, Loraine G. (2003). "Activating Culture through Persuasion Appeals: An Examination of the Bicultural Consumer," *Journal of Consumer Research*, 13 (3), 301-315.

Lee, Cho and Robert T. Green (1991). "Cross-cultural Examination of the Fishbein Behavioral Intentions Model," *International Journal of Business Studies*, 22 (2), 289-305.

Lee, Wei-Na (1993). "Acculturation and Advertising Strategies: A Cross-cultural Study of Chinese and Americans," *Psychology and Marketing*, 10 (5), 381-397.

Lichtenstein, Donald R., Nancy M. Ridgway and Richard G. Netemeyer (1993). "Price Perceptions and Consumer Shopping Behavior: A Field Study," *Journal of Marketing Research*, 30 (2), 234-245.

Locke, Kenneth D. and Kyoung-Deok Baik (2009). "Does Acquiescent Response Style Explain why Koreans are Less Consistent than Americans," *Journal of Cross-Cultural Psychology*, 40 (2), 319-323.

Luedicke, Marius K. (2015). "Indigenes' Responses to Immigrants' Consumer Acculturation: A Relational Configuration Analysis," *Journal of Consumer Research*, 42 (1), 109-129.

Luna, David, Torsten Ringberg, and Laura A. Perracchio (2008). "One Individual, Two Identities: Frame Switching among Biculturals," *Journal of Consumer Research*, 35 (2), 279-293.

Marín, Gerardo, Rymond J. Gamba, and Barbara V. Marín (1992). "Extreme Response Style and Acquiescence among Hispanics," *Journal of Cross-Cultural Psychology*, 23 (4), 498-509.

McCrae, Robert R. (2002). "NEO-PI-R Data from 36 Cultures. Further Intercultural Comparisons," in R. R. McCrae, J. Allik (Eds.), *The five-factor model of personality across cultures*, New York et al.: Kluwer, 105-125.

Mwamwenda, Tuntufye S. (1993). "A Comparison of Two Samples, South Africans and Canadians, on Social Desirability," *Psychological Reports*, 72 (3), 965-966.

Ownbey, Shiretta F. and Patricia E. Horridge (1997). "Acculturation Levels and Shopping Orientations of Asian-American Consumers," *Psychology and Marketing*, 14 (1), 1-18.

Paulhus, D. L. (1991). "Measurement and Control of Response Bias," in J. P. Robinson, P. R. Shaver, L. S. Wrightsman (Eds.), *Measures of Personality and Social Psychological Attitudes*, San Diego: Academic Press, 17-59.

Paulhus, Delroy L. and Oliver P. John (1998). "Egoistic and Moralistic Biases in Self-perception: The Interplay of Self-deceptive Styles with Basic Trait and Motives," *Journal of Personality*, 66 (6), 1025-1060.

Peñaloza, Lisa N. (1989). "Immigrant Consumer Acculturation," *Advances in Consumer Research*, 6, 110-118.

Peñaloza, Lisa N. (1994). "Atravesando Fronteras/Border Crossings: A Critical Ethnographic Exploration of the Consumer Acculturation of Mexican Immigrants," *Journal of Consumer Research*, 21 (1), 32-54.

Peng, Kaiping and Richard E. Nisbett (1999). "Culture, Dialectics, and Reasoning about Contradictions," *American Psychologist*, 54 (9), 741-754.

Peterson, Robert A., Pablo Rhi-Perez, and Gerald Albaum (2014). "A Cross-national Comparison of Extreme Response Style Measures," *International Journal of Market Research*, 56 (1), 89-110.

Perera, Mahendra and Sybil B. G. Eysenck (1984). "A Cross-cultural Study of Personality. Sri Lanka and England," *Journal of Cross-Cultural Psychology*, 15 (3), 353-371.

Psychology and Marketing (2016). *Aims and scope* [available at http://onlinelibrary. wiley.com/journal/10.1002/(ISSN)1520-6793/homepage/ProductInformation.html; accessed at December 16, 2016].

Riordan, Christine M. and Robert J. Vandenberg (1994). "A Central Question in Cross-Cultural Research: Do Employees of Different Cultures Interpret Work-related Measures in an Equivalent Manner?," *Journal of Management*, 20 (3), 643-671.

Ronen, Simcha and Oded Shenkar (2013). "Mapping World Cultures: Cluster Formation, Sources and Implications," *Journal of International Business Studies*, 44 (9), 867-897.

Ross, Catherine E. and John Mirowsky (1984). "Socially-desirable Response and Acquiescence in Cross-cultural Survey of Mental Health," *Journal of Health and Social Behavior*, 25 (2), 189-197.

Schwarz, Norbert (2003). "Self-reports in Consumer Research: The Challenge of Comparing Cohorts and Culture," *Journal of Consumer Research*, 29 (4), 588-594.

Schwartz, Shalom H. (1992). "Universals in the Content and Structure of Values: Theoretical Advances and Empirical Tests in 20 Countries," *Advances in Experimental Social Psychology*, 25, 1-65.

Shiomi, Kunio and Robert Loo (1999). "Cross-cultural Response Styles on the Kriton Adaption-innovation Inventory," *Social Behavior and Personality*, 27 (4), 413-420.

Smith, Peter B. (2004). "Acquiescent Response Bias as an Aspect of Cultural Communication Style," *Journal of Cross-Cultural Psychology*, 35 (1), 50-61.

Statistisches Bundesamt. *Migration und Integration* [available at https://www.destatis.de/EN/FactsFigures/SocietyState/Population/MigrationIntegration/MigrationIntegration.html; accessed December 16, 2016].

Steenkamp, Jan-Benedict E. M. and Hans Baumgartner (1998). "Assessing Measurement Invariance in Cross-national Consumer Research," *Journal of Consumer Research*, 25 (1), 78-90.

Steenkamp, Jan-Benedict E. M., Martijn de Jong, and Hans Baumgartner (2010). "Socially Desirable Response Tendencies in Survey Research," *Journal of Marketing Research*, 47 (2), 199-214.

Stening, Bruce W. and James E. Everett (1984). "Response Styles in a Cross-cultural Managerial Study," *Journal of Social* Psychology, 122 (2), 151-156.

Strizhakova, Yuliya and Robin A. Coulter (2013). "The "Green" Side of Materialism in Emerging BRIC and Developed Markets: The Moderating Role of Global Cultural Identity," *International Journal of Research in Marketing*, 30 (1), 60-82.

Torres, Anna, Tammo H. A. Bijmolt, Josep A. Tribó, and Peter Verhoef (2012). "Generating Global Brand Equity through Corporate Social Responsibility to Key Stakeholders," *International Journal of Research in Marketing*, 29 (1), 13-24.

Tourangeau, Roger T., Lance J. Rips, and Kenneth Rasinski (2009). *The psychology of survey research*. Cambridge: Cambridge University Press.

Üstüner, Tuba and Douglas B. Holt (2007). "Dominated Consumer Acculturation: The Social Construction of Post Migrant Women's Consumer Identity Projects in a Turkish Squatter," *Journal of Consumer Research*, 34 (1), 41-56.

Van Hemert, Dianne A., Fons J. R. van de Vijver, Ype H. Poortinga, and James Georgas (2002). "Structural and Functional Equivalence of the Eysenck Personality Questionnaire

within and Between Countries," *Personality and Individual Differences*, 33 (8), 1229-1249.

Van Herk, Heser, Ype H. Poortinga, and Theo M. Verhallen (2004). "Response Styles in Rating Scales. Evidence of Method Bias in Data from Six EU Countries," *Journal of Cross-Cultural Psychology*, 35 (3), 346-360.

Van Rosmalen, Joost, Hester van Herk, and Patrick J. F. Groenen (2010). "Identifying Response Styles: A Latent-class Bilinear Multinomial Logit Model," *Journal of Marketing Research*, 47 (1), 157-172.

Voss, Kevin E, Donald E. Stem, Lester W. Johnson, and Constantino Arce (1996). "An Exploration of the Comparability of Semantic Adjectives in Three Languages. A Magnitude Estimation Approach," *International Marketing Review*, 13 (5), 44-58.

Wallendorf, Melanie and Michael D. Reilly (1983). "Ethnic Migration, Assimilation, and Consumption," *Journal of Consumer Research*, 10 (3), 292-301.

Watkins, David and Steven Cheung (1995). "Culture, Gender, and Response Bias. An Analysis to the Self-description Questionnaire," *Journal of Cross-Cultural Psychology*, 26 (5), 490-504.

Webster, Cynthia (1996). "Hispanic and Anglo Interviewer and Respondent Ethnicity and Gender: The Impact on Survey Response Quality," *Journal of Marketing Research*, 33 (1), 62-72.

Weech-Maldonado, Robert, Marc N. Elliott, Adetokunbo Oluwole, K. Cameron Schiller, and Ron D. Hays (2008). "Survey Response Style and Differential Use of CAHPS Rating Scales by Hispanics," *Medical Care*, 46 (9), 963-968.

Wells, William D. (1961). "The Influence of Yeasaying on Response Style," *Journal of Advertising Research*, 1 (4), 1-12.

Weijters, Bert and Hans Baumgartner (2012). "Misresponse to Reversed and Negated Items in Surveys: A Review," *Journal of Marketing Research*, 49 (5), 737-747.

Weijters, Bert, Elke Cabooter, and Niels Schillewaert (2010). "The Effect of Rating Scale Format on Response Styles: The Number of Answer Categories and Response Category Labels," *International Journal of Research in Marketing*, 27 (3), 236-247.

Weijters, Bert, Maggie Geuens, and Hans Baumgartner (2013). "The Effect of Familiarity with Response Category Labels on Item Response to Likert Scales," *Journal of Consumer Research*, 40 (2), 368-381.

Will, Edward E. and Jo A. Verdin (1978). "Social Desirability Response Bias in Children's Locus of Control Reports," *Psychological Reports*, 43 (3), 924-926.

Wong, Nancy, Aric Rindfleisch, and James E. Burroughs (2003). "Do Reversed-coded Items Confound Measures in Cross-cultural Consumer Research? The Case of the Material Value Scale," *Journal of Consumer Research*, 30 (1), 72-91.

Wright, George N., Lawrence D. Phillips, Peter C. Whalley, Gery T. Choo, Kee-Ong Ng, Irene Tan, and Aylene Wisudha (1978). "Cultural Differences in Probabilistic Thinking," *Journal of Cross-Cultural Psychology*, 9 (3), 285-299.

Yik, S. M. Michelle, Michael H. Bond, and Delroy L. Paulhus (1998). "Do Chinese Self-enhance or Self-efface? It's a Matter of Domain," *Personality and Social Psychology Bulletin*, 24 (4), 399-406.

Zax, Melvin and Shigeo Takahashi (1967). "Cultural Influences on Response Style: Comparisons of Japanese and American College Students," *Journal of Social Psychology*, 71 (1), 3-10.

APPENDIX PROJECT 4

Appendix A: Overview of Empirical Studies Reviewed

Table A1: Overview of Empirical Studies Reviewed

Study	Sample[1]	Research Field	Response Styles Covered
Abe and Zane (1990)	Student sample; foreign-born Asian American (46), US-born Asian American (29), white American (61)	Personality	SDR
Aday, Chiu, and Andersen (1980)	Non-student sample; Hispanic (1,092), other American (6,695)	Health Care	ARS
Bachman and O'Malley (1984)	Student sample; African and white American (ca. 17,000)	Education	ARS, DARS, ERS
Brengelmann (1959)	Non-student sample; Germany (200), UK (165)	Personality	ARS
Carr and Krause (1978)	Non-student sample; African American (49), Mexican American (11), non-Hispanic white American (106), Puerto Rican (47), other American (6)	Clinical Psychology, Personality	ARS, SDR[2]
Chen, Lee, and Stevenson (1995)	Student sample; Canada (687), China (1,357), Japan (944), USA (2,147)	Education, Psychology	ERS, MRS

Table A1: Overview of Empirical Studies Reviewed (Cont'd)

Study	Sample[1]	Research Field	Response Styles Covered
Chun, Campbell, and Yoo (1974)	Student sample; South Korea (204), USA (187)	Personality	ERS, RR
Clarke (2000)	Sample 1: student sample; Hispanic (146), non-Hispanic Whites (146) Sample 2: student sample; African American (115), non-African American (157) Sample 3: student sample; Australia (121), France (133)	Personality	ERS
Consedine et al. (2012)	Non-student sample; African American (295), American of English Caribbean descent (299), American of Haitian descent (305), American of Dominican descent (160), American of Eastern European descent (151), European American (154)	Health Care	SDR
Crandall, Crandall, and Katkovsky (1965)	Children/student sample; African American (100), white American (856)	Personality	SDR
Crosby (1969)	Non-student sample; English-speaking Canadians (826), French-speaking Canadians (239)	Consumer Behavior	ARS
Davis, Resnicow, and Couper (2011)	Non-student sample; Hispanic and non-Hispanic American (288)	Sociology	ARS, ERS
De Jong et al. (2008)	Non-student sample; Argentina, Austria, Belgium, Brazil, China, Czech Republic, Denmark, France, Germany, Hungary, Ireland, Italy, Japan, Netherlands, Norway, Poland, Portugal, Romania, Russia, Slovakia, Spain, Switzerland, Taiwan, Thailand, UK, USA (between 335 and 1,181 respondents per country, 12,506 respondents in total)	Consumer Behavior	ERS

Table A1: Overview of Empirical Studies Reviewed (Cont'd)

Study	Sample[1]	Research Field	Response Styles Covered
De Langhe et al. (2011)	Study 1: student sample; Dutch-English-French trilinguals; native Dutch speakers (61), native French speakers (59) Study 2-9: student sample; Dutch-English bilinguals (61)	Consumer Behavior	ERS
Dohrenwend (1966)	Non-student sample; African American (27) Irish American (27), Jewish American (27), Puerto Rican (27)	Clinical Psychology	ARS, SDR
Goldsmith, White, and Stith (1987)	Mixed sample; African American (142), White American (166)	Personality	ARS, SDR
Grimm and Chruch (1999)	Study 1: student sample; Philippines (387), USA (610) Study 2: student Sample; Philippines (656), USA (660) Study 3: student Sample; Philippines (518), USA (372)	Personality	ARS, ERS, SDR
Harzing (2006)	Student sample; Austria (53), Brazil (72), Bulgaria (78), Chile (53), China (50), Denmark (44), Finland (87), France (42), Germany (50), Greece (58), Hong Kong (54), India (50), Japan (45), Lithuania (57), Malaysia (65), Mexico (50), Netherlands (109), Poland (54), Portugal (76), Russia (44), Spain (83), Sweden (62), Taiwan (60), Turkey (78), UK (46), USA (61)	Personality	ARS, DARS, ERS, MRS
Heine and Lehman (1995)	Student sample; Canada (74), Japan (93)	Personality	SDR

Table A1: Overview of Empirical Studies Reviewed (Cont'd)

Study	Sample[1]	Research Field	Response Styles Covered
Hofstede (2001)	Non-student sample; Argentina (1,145), Australia (1,919), Austria (1,247), Bahamas (8), Belgium (2,385), Bolivia (19), Brazil (2,385), Canada (3,576), Chile (164), Colombia (427), Costa Rica (75), Denmark (1,304), Dominican Republic (20), Ecuador (73), Finland (802), France (11,337), West-Germany (11,384), Greece (238), Guatemala (99), Hong Kong (88), Honduras (25), Indonesia (91), India (231), Iran (231), Ireland (251), Iraq, Israel (357), Italy (1,797), Jamaica (73), Japan (6,448), South Korea (56), Malaysia (72), Mexico (1,016), Netherlands Antilles (16), Netherlands (1,797), Nicaragua (20), Norway (819), New Zealand (413), Pakistan (107), Panama (81), Peru (290), Philippines (319), Portugal (243), South Africa (867), Salvador (70), Singapore (58), Spain (1,802), Sweden (2,432), Switzerland (2,111), Taiwan (71), Thailand (80), Trinidad (22), Turkey (168), Uruguay (136), UK (6,967), USA (3,967), Venezuela (535), South Vietnam (24), Yugoslavia (248), Arabic-speaking region: Iraq, Kuwait, Lebanon, Libya, Saudi Arabia, United Arab Republic (141), East Africa: Egypt, Ethiopia, Kenya, Tanzania, Zambia (46), West Africa: Ghana, Nigeria, Sierra Leone (43)	Personality	ARS
Hui and Triandis (1989)	Non-student sample; Hispanic American (59), non-Hispanic American (60)	Human Ressource Management, Psychology	ERS
Javeline (1999)	Non-student sample; Kazakhs Russian, other nationals living in Kazakhstan (1,986)	Politics	ARS

Table A1: Overview of Empirical Studies Reviewed (Cont'd)

Study	Sample[1]	Research Field	Response Styles Covered
Johnson et al. (2005)	Non-student sample; Australia (459), Belgium (995), Brazil (1,092), Czech Republic (749), France (1,028), Germany (1,751), Hong Kong (502), Hungary (750), India (1,525), Italy (1,054), Japan (726), Malaysia (500), Mexico (1,420), Philippines (508), Poland (788), Portugal (591), Singapore (519), Turkey (1,000), UK (4,313)	Human Ressource Management	ARS, ERS
Johnson et al. (1997)[3]	Non-student sample; African American (111), Mexican American (112), Puerto Rican (92), non-Hispanic white American (108)	Health Care	ARS, ERS, SDR
Lalwani, Shavitt, and Johnson (2006)	Study 1/sample 1: student sample; Singapore (30), USA (59) Study 1/sample 2: student sample; Korean-American (31), other American (33) Study 2: student sample; American of mixed descendent (124)	Personality	SDR
Lee and Green (1991)	Student sample; South Korea (217), USA (212)	Consumer Behavior	ERS, MRS
Locke and Baik (2009)	Student sample; South Korea (55), USA (62)	Personality	ARS

Table A1: Overview of Empirical Studies Reviewed (Cont'd)

Study	Sample[1]	Research Field	Response Styles Covered
Marín, Gamba, and Marín (1992)	Sample 1: non-student sample; Hispanic American (263), non-Hispanic white American (150) Sample 2: non-student sample; Hispanic American (363), non-Hispanic white American (229) Sample 3: non-student sample; Hispanic American (245), non-Hispanic white American (243) Sample 4: non-student sample; Hispanic American (1,037), non-Hispanic white American (13,803)	Health Care	ARS, ERS
McCrae (2002)	Mixed sample; Austria (444), Belgium (1,119), Canada (848), China (201), Croatia (722), Czech Republic (570), Denmark (1,213), Estonia (1,037), France (1,066), Germany (3,730), Hong Kong (122), Hungary (312), India (473), Indonesia (172), Italy (690), Japan (681), Malaysia (451), Netherlands (1,305), Norway (1,142), Peru (439), Philippines (897), Portugal (1,880), Russia (510), South Arica (274), South Korea (2,946), Spain (196), Sweden (720), Switzerland (107), Taiwan (544), Turkey (260), USA (1,462), Yugoslavia (1,120), Zimbabwe (312)	Personality	RR
Mwamwenda (1993)	Student sample; Canada (190), South Africa (86)	Personality	SDR
Perera and Eysenck (1984)	Non-student sample; Sri Lanka (1027), secondary data for UK	Personality	SDR

Table A1: Overview of Empirical Studies Reviewed (Cont'd)

Study	Sample[1]	Research Field	Response Styles Covered
Peterson, Rhi-Perez, and Albaum (2014)	Student sample; Argentina (59), Australia (57), Austria (50), Belgium (52), Bolivia (71), Brazil (131), Canada (128), Chile (54), China (30), Columbia (149), Denmark (75), France (150), Germany (242), Greece (49), Honduras (39), Hong Kong (113), Hungary (26), Iceland (46), Ireland (38), Malta (49), Mexico (72), Morocco (109), Netherlands (47), New Zealand (88), Norway (183), Philippines (43), Senegal (109), Singapore (117), South Korea (86), Spain (174), Thailand (52), Tunisia (212), Turkey (95), UK (148), USA (2,949), Vietnam (54)	Business Ethics	ERS, RR
Riordan, Vandenberg (1994)	Non-student sample; Korea (195), USA (162)	Human Ressource Management	ARS
Ross and Mirowsky (1984)	Non-student sample; Hispanic and non-Hispanic White American (330), Mexican (133)	Clinical Psychology	ARS, SDR
Shiomi and Loo (1999)	Student sample; Canada (144), Japan (144)	Personality	ERS, MRS
Smith (2004)	Use of data by Hofstede (1980; 2001), House et al. (1999), House, Hanges, Javidan, Dorfman, and Gupta (2004), Leung et al. (2002), Schwartz (1994), Smith et al. (2002), Smith, Dugan, and Trompenaars (1996), Trompenaars (1993), van Hemert, van de Vijver, and Poortinga (2002)	Personality	ARS

Table A1: Overview of Empirical Studies Reviewed (Cont'd)

Study	Sample[1]	Research Field	Response Styles Covered
Steenkamp, de Jong, and Baumgartner (2010)	Non-student sample; Argentina (396), Austria (392), Belgium (508), Brazil (394), China (412), Czech Republic (488), Denmark (517), France (402), Germany (638), Hungary (576), Ireland (548), Italy (397), Netherlands (503), Norway (547), Poland (391), Portugal (434), Romania (431), Russia (389), Slovakia (378), Spain (544), Sweden (418), Switzerland (393), Thailand (402), Ukraine (355), UK (355), USA (1,181)	Consumer Behavior, Personality	ERT, MRT
Steenkamp and Baumgartner (1998)	Non-student sample; Belgium (990), Greece (1,153), UK (974)	Consumer Behavior	ARS
Stening and Everett (1984)	Manager sample; Hong Kong (95), Indonesia (112), Japan (769), Malaysia (113), Philippines (111), Singapore (190), Thailand (95), UK (128), USA (34)	General Management	ERS, MRS
Van Hemert et al. (2002)	Non-student sample; Australia (1,452), Bangladesh (1,075), Brazil (1,396), Bulgaria (1,038), Canada (1,652), Chile (67), China (2,097), Czechoslovakia (804), Egypt (1,330), Finland (949), France (866), Greece (1,301), Hong Kong (732), Hungary (962), Iceland (1,144), India (2,275), Iran (624), Ireland (2,804), Israel (2,412), Italy (2,609), Japan (258), Netherlands (1,401), Nigeria (430), Norway (802), Poland (120), Puerto Rico (1,094), Russia (1,067), Saudi-Arabia (600), Singapore (994), Spain (2,986), Sri Lanka (1,027), Sweden (126), Uganda (1,476), UK (17,725), USA (4,153), West-Germany (2,538), Yugoslavia (1,430), Zimbabwe (2,758)[4]	Personality	SDR

Table A1: Overview of Empirical Studies Reviewed (Cont'd)

Study	Sample[1]	Research Field	Response Styles Covered
Van Herk, Poortinga, and Verhallen (2004)	Sample 1: non-student; France (532), Greece (580), Italy (598) Sample 2: non-student sample; France (313), Germany (295), Greece (281), Italy (299), Spain (303), UK (298) Sample 3: France (550), Germany (634), Italy (806), Spain (453), UK (580)	Consumer Behavior	ARS, ERS
Van Rosmalen, van Herk, and Groenen (2010)	Non-student sample; France (936), Germany (965), Italy (994), Spain (771), USA (848)	Consumer Behavior	ARS, DARS, MRS
Voss et al. (1996)	Student sample; English-speaking (30), Japanese-speaking (30), Putonghua-speaking (30)	Consumer Behavior	
Watkins and Cheung (1995)	Children sample; Australia (497), China (232), Nepal (402), Nigeria (393), Philipines (193)	Personality	ARS, DARS, NCR, RR
Webster (1996)	Non-student sample; Hispanic American (40), non-Hispanic White American (39)	Personality	SDR
Weech-Maldonado et al. (2008)	Non-student sample; Hispanic American (21,381), non-Hispanic White American (231,240)	Health Care	ARS, DARS, ERS
Weijters, Cabooter, and Schillewaert (2010)	Study 1: Netherlands (1,207) Study 2: UK (226)	Consumer Behavior	ARS, DARS, ERS, MRS

Table A1: Overview of Empirical Studies Reviewed (Cont'd)

Study	Sample[1]	Research Field	Response Styles Covered
Weiters, Geuens, and Baumgartner (2013)	Study 1: non-student sample; Netherlands (218) Study 2: non-student sample; French-speaking: Canada (199), France (226), English-speaking: Canada (193), UK (182), USA (182) Study 3: non-student sample; Dutch-speaking: Belgium (644), Netherlands (1,046), English-speaking: UK (908), French-speaking: Belgium (371), France (1,000), Switzerland (303), German-speaking: Germany (993), Switzerland (606), Hungarian-speaking: Hungary (1,003), Italian-speaking: Italy (939), Switzerland (50), Polish-speaking: Poland (802), Romanian-speaking: Romania (970), Slovakian-speaking: Slovakia (1,063), Spanish-speaking: Spain (934), Swedish-speaking: Sweden (974), Turkish-speaking: turkey (914)	Psychology, Consumer Behavior	ERS
Will and Verdin (1978)	Children/student sample; African American (104), White American (118)	Personality	SDR
Wong, Rindfleisch, and Burroughs (2003)	Study 1: non-student sample; Japan (105), Korea (119), Singapore (200), Thailand (200), USA (200) Study 2: non-Student Sample; Japan (116), Thailand (150), USA (126)	Consumer Behavior	ARS
Yik, Bond, and Paulhus (1998)	Student sample; China (130), North America (175)	Personality	SDR

Table A1: Overview of Empirical Studies Reviewed (Cont'd)

Study	Sample[1]	Research Field	Response Styles Covered
Zax and Takahashi (1967)	Student sample; Japan (80), USA (80)	Personality, Psychology	ERS, MRS

1) sample sizes in brackets; 2) Carr and Krause (1978) do not compare SDR across cultural groups; 3) A strong overlap exists between this study and another study by Warnecke et al. (1997). Importantly, both studies use the same sample and report basically the same results. Hence, the existing study only reviews the study of Johnson et al. (1997) to avoid reporting the same results two times; 4) The study by van Hemert et al. (2002) uses secondary data from several studies. In particular, they use the data of Perera and Eysenck (1984).

REFERENCES APPENDIX PROJECT 4

Abe, Jennifer S. and Nolan W. S. Zane (1990). "Psychological Maladjustment among Asian and White American College Students: Controlling for Confounds," *Journal of Counseling Psychology*, 37 (4), 437-444.

Aday, Lu A., Grace Y. Chiu, and Ronald Andersen (1980). "Methodological Issues in Health care Surveys of the Spanish Heritage Population," *American Journal of Public Health*, 70 (4), 367-374.

Bachman, Jerald G. and Patrick M. O'Malley (1984). "Yea-saying, Nay-saying, and Going to the Extremes: Black-white Differences in Response Styles," *Public Opinion Quarterly*, 48 (2), 491-509.

Brengelmann, Johannes C. (1959). "Differences in Questionnaire Responses between English and German Nationals," *Acta Psychologica*, 16, 339-355.

Carr, Leslie G. and Neal Krause (1978). "Social Status, Psychiatric Symptomatology, and Response Bias," *Journal of Health and Social Behavior*, 19 (1), 86-91.

Chen, Chuansheng, Shin-Ying Lee, and Harold W. Stevenson (1995). "Response Style and Cross-cultural Comparisons of Rating Scales among East Asian and North American Students," *Psychological Science*, 6 (3), 171-175.

Chun, Ki-Taek, John B. Campbell, and Jong H. Yoo (1974). "Extreme Response Style in Cross-cultural Research. A Reminder," *Journal of Cross-Cultural Psychology*, 5 (4), 465-480.

Clarke, Irvine III (2000). "Extreme Response Style in Cross-cultural Research: An Empirical Investigation," *Journal of Social Behavior and Personality*, 15 (1), 137-152.

Consedine, Nathan S., Carol Magai, David Horton, and William M. Brown (2012). "The Affective Paradox: An Emotion Regulatory Account of Ethnic Differences in Self-reported Anger," *Journal of Cross-Cultural Psychology*, 43 (5), 723-741.

Crandall, Virginia C., Vaughan J. Crandall, and Walter Katovsky (1965). "A Children's Social Desirability Questionnaire," *Journal of Consulting Psychology*, 29 (1), 27-36.

Crosby, Richard W. (1969). "Attitude Measurement in a Bilingual Culture," *Journal of Marketing Research*, 6 (4), 421-426.

Davis, Rachel E., Ken Resnicow, and Mick P. Couper (2011). "Survey Response, Acculturation, and Culture among a Sample of American Adults," *Journal of Cross-Cultural Psychology*, 42 (7), 1219-1236.

De Jong, Martijn G., Jan-Benedict E. M. Steenkamp, Jean-Paul Fox, and Hans Baumgartner (2008). "Using Item Response Theory to Measure Extreme Response Style in Marketing: A Global Investigation," *Journal of Marketing Research*, 45 (1), 104-115.

De Jong, Martijn G., Rik Pieters, and Jean-Paul Fox (2010). "Reducing Social Desirability Bias through Item Randomized Response: An Application to Measure Underreported Desires," *Journal of Marketing*, 47 (1), 14-27.

De Langhe, Bart, Stefano Puntoni, Daniel Fernandes, and Stijn M. J. van Ossealaer (2011). "The Anchor Contraction Effect in International Marketing Research," *Journal of Marketing Research*, 48 (2), 366-380.

Dohrenwend, Bruce P. (1966). "Social Status and Psychological Disorder: An Issue of Substance and an Issue of Method," *American Sociological Review*, 31 (1), 14-34.

Goldsmith, Ronald E., J. Dennis White, and Melvin T. Stith (1987). "Values of Middle-class Blacks and White: A Replication and Extension," *Psychology and Marketing*, 4 (2), 135-144.

Grimm, Stephanie D. and Timothy A. Church (1999). "A Cross-cultural Study of Response Biases in Personality Measures," *Journal of Research in Personality*, 33 (4), 415-441.

Harzing, Anne-Wil (2006). "Response Styles in Cross-national Survey Research. A 26-Country Study," *International Journal of Cross Cultural Management*, 6 (2), 243-266.

Heine, Steven J. and Darrin R. Lehman(1995). "Social desirability among Canadian and Japanese students," *Journal of Social Psychology*, 135 (6), 777-779.

Hofstede, Geert (1980). *Culture's Consequences*, Beverly Hills et al.: Sage.

Hofstede, Geert (2001). *Culture's Consequences*, Thousand Oaks et al.: Sage.

House, Robert J., Paul J. Hanges, Mansour Javidan, Peter W. Dorfman, and Vipin Gupta (2004). *Culture Leadership, and Organizations: The GLOBE Study of 62 Societies* Thousand Oaks et al.: Sage.

House, Robert J., Paul J. Hanges, S. Aantonio Ruiz-Quintanilla, Peter W. Dorfman, Mansour Javidan, Marcus Dickson, Vipin Gupta, Ikhlas A. Abdalla, Babadjide S. Adetoun, Ram N. Aditya, Hafid Agourram, Adebowale Akande, Bolanle E. Akande, Staffan Akerblom, Carlos Altschul, Eden Alvrez-Backus, Julian Andrews, Maria E. Arias, Mirian S. Arif, Neal M. Askhanasy, Arben Asllani, Guiseppe Audia, Guyula Bakacs, Helena Bendova, David Beveridge, Rabi S. Bhagat, Alejandro Blacutt, Jiming Bao, Domenico Bodega, Muzaffer Bodur, Simon Booth, Annie E. Booysen, Dimitrios Bourantas, Klas Brenk, Felix Brodbeck, Dale E. Carl, Philippe Castel, Chiel-Chen Chang, Sandy Chau, Fenda Cheung, Jagdeep Chhokar, Jimmy Chiu, Peter Cosgriff, Ali Dalmalchian, Jose A. Dela Coleta, Maria F. Dela Coleta, Deanne N. den Hartog, Marc Deneire, Gemma Donnely-Cox, Christopher Earley, Mahmoud A. E. Elgamal, Miriam Erez, Sarah Falkus, Mark Fearing, Richard H. G. Field, Carol Fimmen, Michael Frese, Pingping Fu, Mikhail V. Gratchev, Celia Gurierrez, Mohamed A. Hhashha, Frans M. Hartanto, Markus Hauser, Ingalil Homberg, Marina Holzer, Michael Hoppe, Jon P. Howell, Elena Ibrieva, John C. Ickis, Zakaria Ismail, Slawomir Jarmuz, Jorge C. Jesuino, Li Ji, Kuen-Yung Jone, Geoffrey Jones, Revaz Jorbenadse, Hayat Kabasakal, Mary Keating, Jeffrex C. Kennedy, Jay S. Kim, Giorgi Kipiani, Matthias Kipping, Edvard Konrad, Paul L. Koopman, Fuh-Yeong Juan, Alexandre Kurc, Marie-Francoise Lacassagne, Sang M. Lee, Christopher

Leeds, Francisco Leguizamon, Martin Lindell, Jean Lobell, Fred Luthans, Jerzy Maczyn-ski, Norma Mansor, Gilian Martin, Michael Martin, Sandra M. Martinez, Cecilia McMillen, Emiko Misumi, Jyuji Misumi, Moudi al-Hamoud, Nabil M. Morsi, Phyllisis M. Ngin, Jeremiah O'Connel, Enrique Ogliastri, Nancy Papalexandris, T. K. Peng, Maria M. Preziose, Jose M. Prieto, Boris Rakitsky, Gerhard Reber, Jydeep Roy-Bhattacharya, Amir Rozen, Argio Sabadin, Majhoub Sahaba, Colombia S. de Busamente, Carmen Santana-Melgoza, Daniel A. Sauers, Jette Schramm-Nielsen, Majken Schultz, Zuqi Shi, Camilla Sigfrids, Ahmed Sleem, Kye-Chung Song, Erna Szaboo, Albert C. Teo, Henk Thierry, Jann H. Tjakranegara, Sylvana Trimi, Anne S. Tsui, Pvakanum Ubolwanna, Marius W. van Wyk, Marie Vondrysova, Jürgen Weibler, Celeste Wilderom, Rongxian Wu, Rolf Wunderer, Nik R. N. Yakob, Yongkang Yang, Zuoqiu Yin, Michio Yoshida, and Jian Zhou (1999). "Cultural Influences on Leadership and Organizations: Project GLOBE," in W. H. Mobley M. J. Gessner, V. Arnold (Eds.), *Advances in Global Leadership*, Stamford: JAI Press, 171-233.

Hui, C. Harry, Triandis, Harry C. (1989). "Effects of Culture and Response Format on Extreme Response Style," *Journal of Cross-Cultural Psychology*, 20 (3), 296-309.

Javeline, Debra (1999). "Response Effects in Polite Cultures. A Test of Acquiescence in Kazakhstan," *Public Opinion Quarterly*, 63 (1), 1-28.

Johnson, Timothy, Patrick Kulsea, Young I. Cho, and Sharon Shavitt (2005). "The Relation between Culture and Response Styles. Evidence from 19 Countries," *Journal of Cross-Cultural Society*, 36 (2), 264-277.

Johnson, Timothy, Diane O'Rourke, Noel Chavez, Seymour Sudman, Richard Warnecke, Loretta Lacey, and John Horm (1997). "Social Cognition and Responses to Survey Questions among Culturally Diverse Populations," in L. Lyberg, P. Biemer, M. Collins, E. de Leeuuw, C. Nippo, N. Schwarz, D. Trewin (Eds.), *Survey Measurement and Process Quality*, New York: Wiley, 87-113.

Lalwani, Ashok K., Sharon Shavitt, and Timothy Johnson (2006). "What is the Relation between Cultural Orientation and Socially Desirable Responding?," *Journal of Personality and Social Psychology*, 90 (1), 165-178.

Lee, Cho and Robert T. Green (1991). "Cross-cultural Examination of the Fishbein Behavioral Intentions Model," *International Journal of Business Studies*, 22 (2), 289-305.

Leung, Kwok, Michael H. Bond, Sharon R. de Carrasquel, Carlos Muñoz, Marisela Hernández, Fumio Murakami, Susumu Yamaguchi, Günter Bierbrauer, and Theordore M. Singelis (2002). "Social Axioms: The Search for Universal Dimensions of General Beliefs about How the World Functions," *Journal of Cross-Cultural Psychology*, 33 (3), 286-302.

Locke, Kenneth D. and Kyoung-Deok Baik (2009). "Does Acquiescent Response Style Explain why Koreans are Less Consistent than Americans," Journal of Cross-Cultural Psychology, 40 (2), 319-323.

Marín, Gerardo, Rymond J. Gamba, and Barbara V. Marín (1992). "Extreme Response Style and Acquiescence among Hispanics," Journal of Cross-Cultural Psychology, 23 (4), 498-509.

McCrae, Robert R. (2002). "NEO-PI-R Data from 36 Cultures. Further Intercultural Comparisons," in R. R. McCrae, J. Allik (Eds.), *The five-factor model of personality across cultures*, New York et al.: Kluwer, 105-125.

Mwamwenda, Tuntufye S. (1993). "A Comparison of Two Samples, South Africans and Canadians, on Social Desirability," *Psychological Reports*, 72 (3), 965-966.

Peterson, Robert A., Pablo Rhi-Perez, and Gerald Albaum (2014). "A Cross-national Comparison of Extreme Response Style Measures," *International Journal of Market Research*, 56 (1), 89-110.

Perera, Mahendra and Sybil B. G. Eysenck (1984). "A Cross-cultural Study of Personality. Sri Lanka and England," *Journal of Cross-Cultural Psychology*, 15 (3), 353-371.

Riordan, Christine M. and Robert J. Vandenberg (1994). "A Central Question in Cross-Cultural Research: Do Employees of Different Cultures Interpret Work-related Measures in an Equivalent Manner?," *Journal of Management*, 20 (3), 643-671.

Ross, Catherine E. and John Mirowsky (1984). "Socially-desirable Response and Acquiescence in Cross-cultural Survey of Mental Health," *Journal of Health and Social Behavior*, 25 (2), 189-197.

Schwartz, Shalom H. (1994). "Beyond Individualism/Collectivism: New Cultural Dimensions of Values," in U. Kim, H. C. Triandis, C. Kagitçibasi, S. C. Choi, G. Yoon (Eds.), *Individualism and Collectivism: Theory, Method and Applications*, Thousand Oaks et al.: Sage, 85-119.

Shiomi, Kunio and Robert Loo (1999). "Cross-cultural Response Styles on the Kriton Adaption-innovation Inventory," *Social Behavior and Personality*, 27 (4), 413-420.

Smith, Peter B. (2004). "Acquiescent Response Bias as an Aspect of Cultural Communication Style," *Journal of Cross-Cultural Psychology*, 35 (1), 50-61.

Smith, Peter B., Shaun Dugan, and Fons Trompenaars (1996). "National Culture and the Values of Organizational Employees," *Journal of Cross-Cultural Psychology*, 27 (2), 231-264.

Smith, Peter B., Mark F. Peterson, Shalom H. Schwartz, Abd H. Ahmad, Debo Akadne, Jon A. Andersen, Sabino Ayestaran, Stephen Bochner, Victor Callan, Carlos Davila, Bjorn Ekelund, Pierre-Henri François, Gert Graversen, Charles Harb, Jorge Jesuino, Aristotle Kantas, Lyudmila Karamushka, Paul Koopman, Kwok Leung, Pavla Kruzela, Sigmar Malvezzi, Andrew Mogaji, Shahrenaz Moratazavi, John Munene, Ken Parry, Betty J. Punnett, Mark Radford, Arja Ropo, Jose Saiz, Grant Savage, Bernadette Setiadi, Ritch, Ritch Sorensen, Erna Szabo, Punyacha Teparakul, Aqeel Tirmizi, Sevda Tsvetanoca, Conrad Viedge, Carolyn Wall, and Vladimir Yanchuk (2002). "Cultural Values, Sources of Guidance, and Their Relevance to Managerial Behavior," *Journal of Cross-Cultural Psychology*, 33 (2), 188-208.

Steenkamp, Jan-Benedict E. M. and Hans Baumgartner (1998). "Assessing Measurement Invariance in Cross-national Consumer Research," *Journal of Consumer Research*, 25 (1), 78-90.

Steenkamp, Jan-Benedict E. M., Martijn de Jong, and Hans Baumgartner (2010). "Socially Desirable Response Tendencies in Survey Research," *Journal of Marketing Research*, 47 (2), 199-214.

Stening, Bruce W. and James E. Everett (1984). "Response Styles in a Cross-cultural Managerial Study," *Journal of Social* Psychology, 122 (2), 151-156.

Trompenaars, Fons (1993). *Riding the Waves of Culture*. London: Economist.

Van Hemert, Dianne A., Fons J. R. van de Vijver, Ype H. Poortinga, and James Georgas (2002). "Structural and Functional Equivalence of the Eysenck Personality Questionnaire within and Between Countries," *Personality and Individual Differences*, 33 (8), 1229-1249.

Van Herk, Heser, Ype H. Poortinga, and Theo M. Verhallen (2004). "Response Styles in Rating Scales. Evidence of Method Bias in Data from Six EU Countries," *Journal of Cross-Cultural Psychology*, 35 (3), 346-360.

Van Rosmalen, Joost, Hester van Herk, and Patrick J. F. Groenen (2010). "Identifying Response Styles: A Latent-class Bilinear Multinomial Logit Model," *Journal of Marketing Research*, 47 (1), 157-172.

Voss, Kevin E, Donald E. Stem, Lester W. Johnson, and Constantino Arce (1996). "An Exploration of the Comparability of Semantic Adjectives in Three Languages. A Magnitude Estimation Approach," *International Marketing Review*, 13 (5), 44-58.

Warnecke, Richard B., Timothy Johnson, Noel Chávez, Seymour Sudman, Diane P. O'Rourke, Loretta Lacey and John Horm (1997). "Improving Question Wording in Surveys of Culturally Diverse Populations," *Annals of Epidemiology*, 7 (5), 334-342.

Watkins, David and Steven Cheung (1995). "Culture, Gender, and Response Bias. An Analysis to the Self-description Questionnaire," *Journal of Cross-Cultural Psychology*, 26 (5), 490-504.

Webster, Cynthia (1996). "Hispanic and Anglo Interviewer and Respondent Ethnicity and Gender: The Impact on Survey Response Quality," *Journal of Marketing Research*, 33 (1), 62-72.

Weech-Maldonado, Robert, Marc N. Elliott, Adetokunbo Oluwole, K. Cameron Schiller, and Ron D. Hays (2008). "Survey Response Style and Differential Use of CAHPS Rating Scales by Hispanics," *Medical Care*, 46 (9), 963-968.

Weijters, Bert, Elke Cabooter, and Niels Schillewaert (2010). "The Effect of Rating Scale Format on Response Styles: The Number of Answer Categories and Response Category Labels," *International Journal of Research in Marketing*, 27 (3), 236-247.

Weijters, Bert, Maggie Geuens, and Hans Baumgartner (2013). "The Effect of Familiarity with Response Category Labels on Item Response to Likert Scales," *Journal of Consumer Research*, 40 (2), 368-381.

Will, Edward E. and Jo A. Verdin (1978). "Social Desirability Response Bias in Children's Locus of Control Reports," *Psychological Reports*, 43 (3), 924-926.

Wong, Nancy, Aric Rindfleisch, and James E. Burroughs (2003). "Do Reversed-coded Items Confound Measures in Cross-cultural Consumer Research? The Case of the Material Value Scale," *Journal of Consumer Research*, 30 (1), 72-91.

Yik, S. M. Michelle, Michael H. Bond, and Delroy L. Paulhus (1998). "Do Chinese Self-enhance or Self-efface? It's a Matter of Domain," *Personality and Social Psychology Bulletin*, 24 (4), 399-406.

Zax, Melvin and Shigeo Takahashi (1967). "Cultural Influences on Response Style: Comparisons of Japanese and American College Students," *Journal of Social Psychology*, 71 (1), 3-10.

CURRICULUM VITAE

Education

04/2012– ongoing	**Doctoral Student** **Chair for Marketing and Market Research (Prof. Dr. Marc Fischer)** **University of Cologne, Germany** Dissertation on the topic "Essays on Consumer Perception - Applications to Inflation, Business-to-Business Brands and Response Styles"
10/2006 – 03/2012	**Diploma Studies in Industrial Engineering, Mechanical Engineering Specialization** **Rheinisch-Westfälisch Technische Hochschule, Aachen** · Degree: Diplom-Wirtschafts-Ingenieur (equivalent to Master) · Majors: Production Engineering
08/1998 – 06/2006	**Stiftisches Gymnasium Düren, Germany** Abitur (equivalent to A level)

Previous Work Experience

04/2012 – ongoing	**Research Assistant** **Chair for Marketing and Market Research (Prof. Dr. Marc Fischer)** **University of Cologne, Germany** · Tutor for the courses "Concepts of Marketing Mix Management" (Bachelor level) and "Methods of Marketing Mix Management" (Bachelor level) · Advisor for numerous seminar, bachelor, master, and diploma theses · Supervisor of a business project for students of the CEMS master in International Management
05/2010 – 03/2012	**Student Assistant** **Laboratory of Machine Tools and Production Engineering** **Rheinisch-Westfälisch Technische Hochschule, Aachen, Germany** · Conception and preparation of workshops and teaching materials · Data collection and analysis · Literature search · Preparation and attendance to industry meetings
11/2010 – 02/2011	**Intern** **SIG Combibloc, Linnich, Germany** · Evaluation of production capacities · Development of a sales guideline · Preparation of consumer research
10/2009 – 02/2010	**Tutor** **Chair for Business Administration and Corporate Finance** **Rheinisch-Westfälisch Technische Hochschule, Aachen, Germany** Tutor for Corporate Finance
03/2007 – 08/2009	**Wardrobe Supervisor and Doorkeeper** **Theater Aachen, Aachen, Germany**
02/2007 – 03/2007	**Technical Intern** **RWE Power AG, Eschweiler, Germany**
06/2006 – 08/2006	**Technical Intern** **Carl KRAFFT and Söhne Gmbh and Co. KG, Düren, Germany**
2005 – 2007	**Counter Service** **Wasserinformationszentrum Eifel, Heimbach, Germany**

Publications

In Refereed Journals

2015	**Drivers of B2B Brand Strength – Insights from an International Study across Industries,** *Schmalenbach Business Review*, Vol. 67 (January), 114-137 (together with Marc Fischer, Manfred Krafft, and Kay Peters)

German Books Chapters

forthcoming	**Treiber von B-to-B-Marken – Ergebnisse einer internationalen branchenübergreifenden Studie (Drivers of B2B Brand Strength – Insights from an International Study across Industries)**, in C. Baumgarth (Ed.): Handbuch B-to-B-Markenführung (together with Marc Fischer and Manfred Krafft)

In German Applied Business Journals

2015	**The Relevance of Brands (Die Relevanz von Marken),** *Markenartikel*, Issue 3, 50-52 (together with Marc Fischer, Jesko Perrey, and Tjark Freundt)

Conference and Symposium Presentations

2016	**Consumers' Perception of Price Inflation,** 38th Marketing Science Conference, Shanghai, China, June 16 – 18
	Consumers' Perception of Price Inflation, 20th Doctoral Colloquium at Simon-Kucher and Partners, Bonn, Germany, June 02-03
2015	**Consumers Perception of Price Inflation,** · 17th Annual Meeting Quantitative Marketing, Münster, Germany, September 20 – 22

Other Activities

2009 – ongoing	**Member of the State Band** Volksmusikerbund Nordrhein-Westfalen, Germany
2001 – ongoing	**Member of the Symphonic Wind Orchestra** Musikschule, Düren, Germany
2004 – 2012	**Member of the Parish Council** Catholic Parish St. Nikolaus in Heimbach, Germany
2002 – 2006	**Youth Leader** Catholic Parish St. Nikolaus in Heimbach, Germany
2002 – 2005	**Swimming Coach** Deutsche Lebens-Rettungs-Gesellschaft, Mechernich, Germany

Language and Software Skills

Languages	**German** (native) **English** (business fluent) **French** (basic) **Spanish** (basic) **Latin**
IT	**Microsoft Office** (very good skills) **SPSS** (very good skills) **LIMDEP** (good skills) **AMOS** (good skills) **HTML** (basic skills) **Java** (basic skills)

Interests

Music (clarinet in several ensembles, arranging music), Social Activities, Traveling, Reading

Köln, den 20. Dezember 2016