Empirical Investigations on User Perception and the Effectiveness of Persuasive Technologies

D I S S E R T A T I O N
of the University of St. Gallen,
School of Management,
Economics, Law, Social Sciences
and International Affairs
to obtain the title of
Doctor Oeconomiae

submitted by

Peter Mayer

from

Germany

Approved on the application of

Prof. Dr. Elgar Fleisch

and

Prof. Dr. Frédéric Thiesse

Dissertation Nr. 4028

Logos Verlag Berlin, 2012

The University of St. Gallen, School of Management, Economics, Law, Social Sciences and International Affairs hereby consents to the printing of the present dissertation, without hereby expressing any opinion on the views herein expressed.

St. Gallen, 21. Mai 2012

The President:

Prof. Dr. Thomas Bieger

For My Dear Family

Acknowledgements

Having reached the final lines of my dissertation, starting a disseration project seems akin to setting off for a long sailing adventure. With your mind being full of plans and expectations of where to go and what to discover, you start off, but you recognize soon that stormy waters turn your plans and expectations into doubts and concerns. Having overcome the first tempests, you gain confidence in your capabilities, and you move on with new courage. After a while, your plans are disturbed again. This time, it's not tempests or stormy waters but lovely places that make you stay much longer than planned. But you remember to pick up your plans and follow your route to reach your home port in time. Being at home, you reflect on what you experienced - about the world and about yourself. And you conclude that it was the hard times where you grew, the lovely places where you became inspired, and - most important - the many people you met, who made your journey a unique experience.During the three years at the Institute of Technology, many outstanding people accompanied my way. To them, I am deeply grateful for the support they had given to me in writing this thesis.

First and foremost, I want to thank Prof. Elgar Fleisch - not only for making this dissertation possible, not only for his scientific supervision, but also for his personal inspiration, which goes far beyond the narrow scope of this piece of work. I also want to thank Prof. Frédéric Thiesse for becoming my co-supervisor, his scientific guidance, and the valuable feedback on my work - always garnished with a pinch of humor.

Furthermore, I want to thank all my colleagues at the Institute of Technology Management for the great collaboration and the pleasant time we had. In particular, I want to thank Dirk Volland, Lasse Stünitz, Tobias Kowatsch, Andreas Budde and Johannes Paefgen for the many discussions we had on the research work presented in this thesis. I am also grateful to Christian Purucker and Christian Hildebrand from the Center for Customer Insight (CCI-HSG) for their feedback on statistical evaluations. A special 'thank you' goes to Edward Ho, who spent so many hours for proof-reading the manuscript. And I want to thank Elisabeth Vetsch-Keller and Jeannette Strompen for their great administrative support, and - much more important - for their contribution to the outstandingly positive atmosphere at ITEM.

Finally I want to express my deepest gratitude towards my family for their unconditional support and their imperturbable trust and confidence, which was an indispensable anchor - in stormy waters as well as at lovely places.

St. Gallen, May 2012 Peter Mayer

Abstract

Persuasive technologies denote technical approaches that are aimed at influencing human behavior. A multitude of persuasive technologies have been evaluated with regard to their technical feasibility, their persuasive effectiveness, and their user acceptance.

The present dissertation extends the existing body of research along three dimensions. These are firstly the user acceptance of persuasive environments; secondly the potential of profiling to increase persuasive effectiveness; and thirdly, the user acceptance of persuasive business models.

A first study investigates the example of a persuasive kitchen environment, focusing on which factors influence the acceptance of such an approach. Study results show that the relatively moderate acceptance level primarily depends on performance considerations and social influence, and only to a lesser degree on effort and usability aspects. Individual perception furthermore varies with gender, predisposition towards the proposed functionality, and general attitude towards new technologies.

A second study analyzes to which degree adapting persuasive strategies to personality traits can increase the effectiveness of persuasive technologies. Experimental results show that such adaptation can increase the effect of persuasive messages in comparison to the application of a single persuasive strategy, but that a random combination of different strategies is equally effective.

In a third study, we investigate the example of behavior-based automobile insurance, focusing on which factors influence the consumer acceptance of such a persuasive business model. The relatively high acceptance level is primarily influenced by the saving potential that is achievable by appropriate driving behavior. Other important factors are the expected effect on driving pleasure, expected technical performance, and the general attitude towards the proposed insurance model. In contrast, trust in the insurance company, privacy of collected data, increased complexity and social influence play only a minor role for an adoption decision. Furthermore, the study has shown that such insurance models hold the potential to increase one's willingness for behavioral change.

Zusammenfassung

Mit persuasiven Technologien bezeichnet man technologische Ansätze, welche eine Verhaltensbeeinflussung zum Ziel haben. Bislang wurden eine Reihe persuasiver Technologien hinsichtlich ihrer technischen Machbarkeit, ihrer verhaltensbeeinflussenden Wirkung, und hinsichtlich ihrer Nutzerakzeptanz untersucht.

Die vorliegende Dissertation erweitert die bestehenden Forschungsergebnisse entlang dreier Dimensionen, nämlich erstens der Akzeptanz persuasiver Umgebungen, zweitens des Potenzials der Effektivitätssteigerung durch Personalisierung, und drittens der Akzeptanz persuasiver Geschäftsmodelle.

In einer ersten Studie wird am Beispiel einer persuasiven Küchenumgebung untersucht, welche Faktoren die Akzeptanz eines solchen Ansatzes beeinflussen. Im Ergebnis zeigt sich, dass das vergleichsweise moderate Akzeptanzniveau vor allem von Nutzenüberlegungen und sozialen Einflüssen abhängt, weniger dagegen von der erwarteten Benutzbarkeit. Wahrnehmungsunterschiede wurden festgestellt hinsichtlich des Geschlechts der Probanden, der Prädisposition bezüglich der dargebotenen Funktionalität, und der generellen Einstellung gegenüber neuen Technologien.

In einer zweiten Studie wird anhand eines Experiments untersucht, in wie weit eine Anpassung an die Charakterzüge eines Menschen die Effektivität persuasiver Technologien erhöhen kann. Es zeigt sich, dass eine solche Personalisierung die Wirkung von persuasiven Nachrichten im Vergleich zur Anwendung einer einzelnen Strategie erhöhen kann, jedoch eine zufällige Kombination verschiedener Strategien ebenso wirkungsvoll ist.

In einer dritten Studie wird am Beispiel einer verhaltensbasierten Autoversicherung untersucht, welche Faktoren die Akzeptanz eines solchen persuasiven Geschäftsmodells beeinflussen. Das vergleichsweise hohe Akzeptanzniveau wird vor allem von der durch angemessenes Fahrverhalten erreichbaren Prämienreduktion beeinflusst. Weitere wichtige Einflussfaktoren sind die erwartete Auswirkung auf die Fahrfreude, die erwartete Verlässlichkeit der Technologie, und die generelle Einstellung zu solchen Versicherungsmodellen. Vertrauen in den Versicherer, vertrauliche Behandlung der erhobenen Daten, gesteigerte Komplexität sowie sozialer Einfluss spielen dagegen eine vergleichsweise geringe Rolle. Ausserdem zeigt die Studie, dass solche Versicherungsmodelle das Potenzial haben, die Bereitschaft zu einer positiven Verhaltensänderung zu erhöhen.

Table of Contents

List of Figures

List of Tables

Abbreviations

α	Cronbach's a
AIC	Akaike Information Criterion
ANOVA	Analysis of Variance
AT	Attitude
AVE	Average Variance Extracted
BI	Behavioral Intention
BIC	Bayes Information Criterion
BIC	Behavioral Intention to Change
CR	Composite Reliability
CS	Cost Sensitivity
DS	Driving Style
EE	Effort Expectancy
EFA	Exploratory Factor Analysis
EUR	Euro
GPS	Global Positioning System
IMP	Importance
IS	Information System
IT	Information Technology
KMO	Kaiser-Meyer-Olkin
LMM	Linear Mixed Models
LRT	Likelihood Ratio Test
LV	Latent Variable
ML	Maximum Likelihood
MSA	Measure of Sampling Adequacy
OLS	Ordinary Least Squares
PCA	Principal Component Analysis
PE	Performance Expectancy
PFC	Preference for Consistency
PIIT	Personal Innovativeness in IT
PJ	Perceived Enjoyment
PLS	Partial Least Squares
PP	Perceived Privacy
PRE	Personal Relevance
ρ	Composite Reliability
RA	Relative Advantage

REML	Restricted Maximum Likelihood
SEM	Structural Equation Modeling
SI	Social Influence
Sig.	Significance
Std. Dev.	Standard Deviation
TAM	Technology Acceptance Model
TPB	Theory of Planned Behavior
TRA	Theory of Reasoned Action
TRA	Trust
USD	US-Dollar
UTAUT	Unified Theory of Acceptance and Use of Technology
VIF	Variance Inflation Factor

"Vastu-sâmye citta-bhedât tayor vibhaktaï panthâï."

People perceive the same object differently, as each person's perception follows a separate path from another's.

(Patañjali)

I Introduction

Over the last decade, a new class of computational systems has emerged, which are designed to induce changes of human attitudes and behaviors (Fogg 1999). Reeves and Nass (1996) paved the ground for so-called *persuasive technologies* by demonstrating that people treat feedback from computers and other media in a similar way as feedback from human-beings. (Fogg 2002) has shown that people respond socially to persuasive attempts from computer systems. Like in human interaction, persuasive attempts are more successful if the "personality" of the computer matches the personality of the user. He has furthermore shown in various experiments that persuasive strategies like praise, reciprocity or authority, which have been widely investigated in social psychology, remain effective when implemented in information systems.

The effectiveness of persuasive technologies has been demonstrated in a large number of implementations, for example to support people in physical activity (Consolvo et al. 2009; Lacroix et al. 2009; Lin et al. 2006), sustaining a healthy diet (Lo et al. 2007), stop smoking (Grolleman et al. 2006), losing weight (Arteaga et al. 2009), saving energy (Loock et al. 2011; Shiraishi et al. 2009), or changing transportation habits (Froehlich et al. 2009; Lee et al. 2010; Meschtscherjakov et al. 2009).

Research on persuasion has a long tradition in social psychology. Researchers identified and investigated a number of persuasive mechanisms such as social norms, authoritarian behavior or the power of little favors, which make people follow the will of others (Cialdini 2008). Information Systems (IS) research has borrowed from these results to increase the effectiveness of persuasive technologies and to derive design guidelines (Fogg and Hreha 2010). Besides the evaluation of technical implementations, research in the IS domain has focused on impact assessments of persuasive systems, and on the investigation of user perception and acceptance.

This dissertation extends the existing body of work along three dimensions. First, we investigate the acceptance and perception of persuasive environments (as opposed to isolated persuasive technologies and applications), which surround and support users in their daily lives. Second, we analyze whether profiling, i.e. the adaptation of persuasive strategies to personality traits, can increase the effectiveness of persuasive technologies. Third, we investigate consumer acceptance of persuasive business models that provide monetary incentives for certain behavior, which is monitored and evaluated by technical means.

I.1 Objectives and Research Questions

The goal of this dissertation is to evaluate whether persuasive environments, profiling, and persuasive business models are appropriate means to increase the acceptance and effectiveness of persuasive technologies. With regard to *persuasive environments*, our objective is to analyze which factors influence their acceptance and perception. Regarding the assessment of *profiling* as an approach to adapt persuasive technologies to the character traits of their users, our objective is to evaluate whether such adaptive persuasive systems are superior to non-adaptive systems. Two objectives are pursued with regard to *persuasive business models*. First, we intend to evaluate which factors influence consumer acceptance of such business models, and second we investigate whether such business models can be expected to lead to a change in behavior.

To summarize, this dissertation is intended to answer the following research question:

Can persuasive environments, profiling, and persuasive business models increase the acceptance and effectiveness of persuasive technologies?

This research question is broken down into three sub-questions, which relate to three concrete case examples analyzed in this dissertation:

Q1: *Which factors influence the acceptance of a persuasive kitchen environment?*

Q2: *Can profiling improve the effectiveness of persuasive messages?*

Q3: *Which factors influence the acceptance of behavior-based automobile insurance and their effect on the willingness to change driving behavior?*

I.2 Research Process and Data Collection

To answer the research questions raised in this dissertation, three quantitative empirical studies have been conducted. In a first study, we investigate how consumers perceive and accept a persuasive kitchen environment, which aims at supporting healthy nutrition habits. The kitchen environment - as an example for persuasive environments - proposes recipes that are attuned to the nutrition goals of the user. It furthermore guides him in the preparation process, prepares appropriate shopping lists, captures new recipes by monitoring their preparation based on sensors embedded in kitchen devices, and allows for assessing nutrition behavior. Applying a technology acceptance research model, the study is aimed at investigating how consumers perceive certain aspects of the proposed kitchen environment, and how certain personality traits influence the perception of the described scenarios.

Closely linked to the application domain of the first study, the second study investigates to which degree an adaptation of persuasive strategies to individual character traits may increase the effectiveness of persuasive technologies. Persuasive environments provide the technical infrastructure to induce a desired behavioral change, but psychological mechanisms have turned out to be effective and even necessary to motivate a person to adopt the intended behavior (Fogg 2002). At critical moments where people decide about a certain behavior - such as selecting a recipe or creating the shopping list in the kitchen example - persuasive principles should be applied to remind people about their self-selected goals. Different persuasive principles have been identified in psychological research (Cialdini 2008).The experiment presented in this dissertation has shown that an adaptation of the applied persuasive principle to individual character traits may increase the effectiveness of persuasive technologies.

The third study investigates a different approach to influence human behavior. In a multitude of domains, an external entity regards a certain behavior as desirable. For example, public authorities and health insurance companies are interested in a healthy way of living of citizens. Similarly, automobile insurance and leasing companies benefit from a careful driving behavior of their customers. Wherever the consequences of negative behavior are externalized, a third party, who bears the associated risk, is interested to motivate appropriate behavior. Behavior-based business models hold the potential to monitor and assess actual behavior, and to reward positive or sanction negative behavioral patterns. Using the example of behavior-based automobile insurance, we investigate to which degree people are willing to accept insurance premiums that depend on driving behavior. It can be expected that the achievable monetary advantage increases acceptance, whereas consequences for the pleasure of driving, the feeling to be monitored or trust in the technology and the insurance company may have a negative impact on the perception of the proposed insurance model. The relative importance of these aspects and the necessary monetary incentives are analyzed in this third study in order to draw conclusions for the potential of behavior-based business models.

In summary, the present dissertation presents novel findings with regard to consumer acceptance and effectiveness of persuasive technologies. The findings are derived from three quantitative empirical studies. From every study, conclusions are drawn for the design of related applications and product offerings, which may help persuasive technologies leave their niche markets and laboratory environments and become integral parts of everyday products and business models whilst supporting people in achieving their individual goals to change certain behavioral patterns.

I.3 Structure of the Thesis

The structure of the present dissertation is summarized in Figure 1. The introductory chapter I motivates the objectives and research questions raised in this thesis, summarizes research process and data collection, and outlines the thesis structure. The following three chapters present three self-contained studies on certain aspects of persuasive technologies. Chapter II investigates how consumers perceive and accept a persuasive kitchen environment. Based on a technology acceptance research model, an empirical survey has been conducted. As evaluation methodology, structural equation modeling (SEM) with partial least squares (PLS) as estimation method has been applied. Chapter III presents an experiment that has been conducted to analyze whether persuasion profiling may increase the effectiveness of persuasive technologies. The experiment has been evaluated by the means of linear mixed models. Chapter IV investigates consumer perception and acceptance of persuasive business models through the example of behavior-based automobile insurance. It furthermore explores whether the proposed insurance model can be expected to achieve the intended behavioral change. Similar to the first study, a technology acceptance research model has been applied, which was evaluated by the means of structural equation modeling and PLS estimation. Chapter V concludes this dissertation by summarizing the findings of the three studies and answering the research questions raised in Chapter I.

Chapters II, III, and IV follow the same structure. Each chapter begins with a description of the example case. Next, context-specific related work is summarized. Then, the research design, i.e. the research model or the experimental design respectively, is explained, followed by a description of the data collection process. Then, the data analysis process and evaluation results are presented. This includes an explanation of the selected evaluation methodology, the actual evaluation, and in the case of chapters II and IV, an exploratory analysis of the survey results by the means of descriptive statistics are presented. Each of the three chapters concludes with a discussion of theoretical and practical implications.

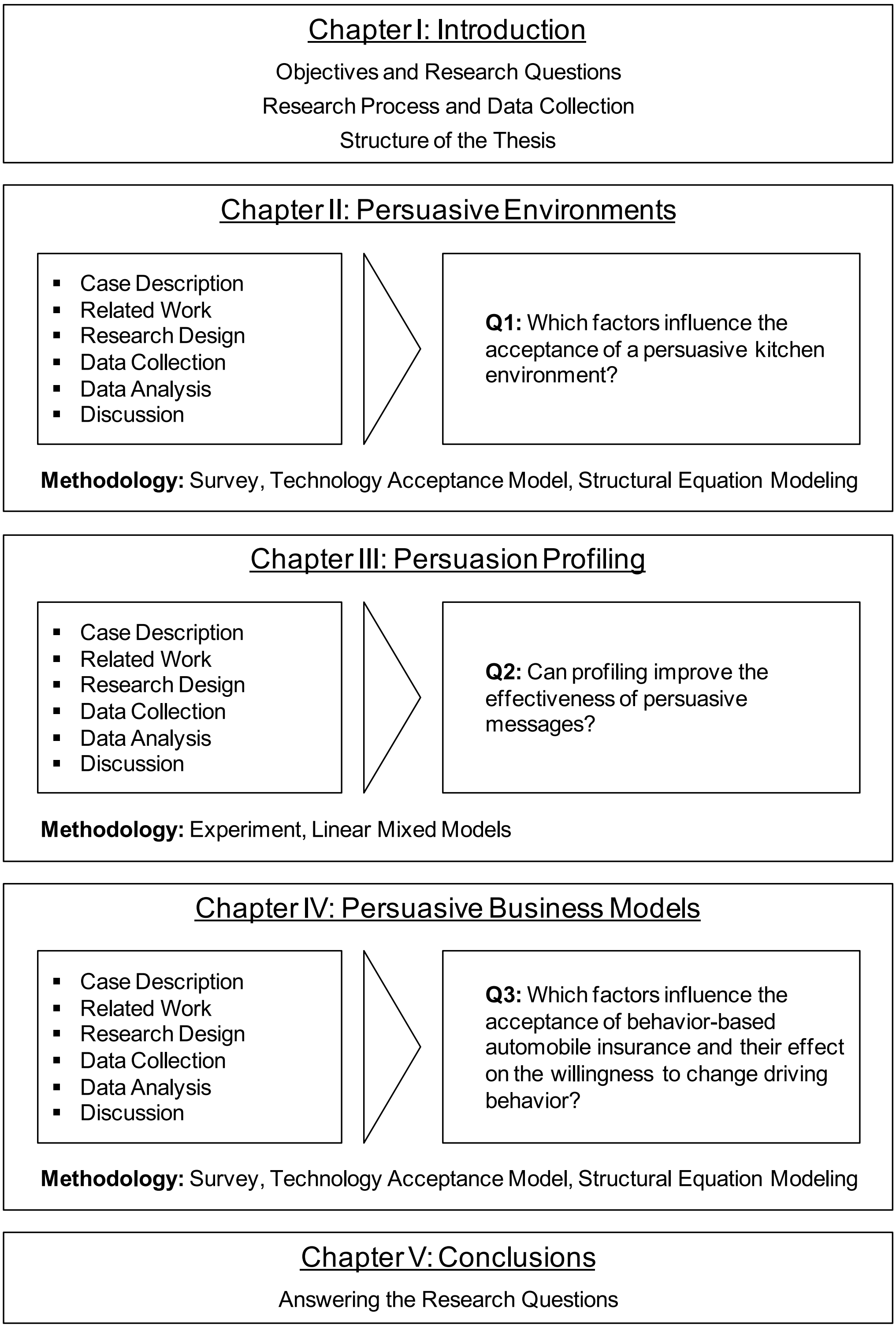


Figure 1: Structure of the Thesis

II Persuasive Environments

II.1 Case Description

The goal of this first case study is to find answers on the first research question raised in this dissertation, namely:

Q1: *Which factors influence the acceptance of a persuasive kitchen environment?*

To investigate this research question, we developed a comprehensive persuasive kitchen scenario that encompasses several household appliances and digital services to support users in preparing meals and improving their nutrition habits. The scenario was developed in close cooperation with our research partner Philips Research to ensure industrial relevance and basic market need.

The focus of our study is on a complex persuasive kitchen environment, which incorporates different kitchen tools that interact with each other and show context-aware behavior. Table 1 summarizes the functional blocks of this environment. The *Cooking Guide* is the central user interface for the persuasive kitchen environment. It provides meal recommendations based on available ingredients and kitchen utensils as well as personal preferences. To guide users in their preparation process, textual and visual presentations provide step-by-step instructions that are synchronized with the actual preparation progress. Persuasive kitchen utensils can be parameterized according to recipe information, and they give feedback on ongoing activities and status information (e.g. temperature, weight, processing times). A recipe memorization function allows for recording preparation processes, including sensor information from the kitchen tools. Once a recipe is chosen, the user can retrieve a shopping list either as a printout or on a mobile phone. The shopping list considers which ingredients are already available in the household. Finally, the user can monitor his or her nutrition habits. Consumption in the persuasive kitchen is automatically recorded, and a mobile application enables users to track non-domestic consumption.

The rationale behind this persuasive kitchen environment is to interact with the user at the points in time that determine the healthiness of a meal. These points in time are the selection of a recipe, shopping of ingredients, and meal preparation. The monitoring function gives feedback to the user in order to motivate him for a sustainable change in nutrition habits.

Table 1: Sub-Scenarios of the Persuasive Kitchen Environment

Sub-Scenario	Description
1. Meal Recommendation	The *Cooking Guide*, a Tablet-PC-like device, proposes recipes based on the ingredients and utensils available in the home and the user's nutritional preferences.
2. Cooking Guidance	The *Cooking Guide* guides the user through the preparation process of a selected dish. While videos and textual descriptions provide step-by-step explanations to the user, kitchen tools continuously communicate with the Cooking Guide and automatically to monitor the preparation process and to configure themselves in accordance with current dish preparation.
3. Recipe Memorization	The persuasive kitchen environment memorizes a user's meal preparation. It tracks the preparation process including ingredients and equipment settings and stores it as a recipe for future retrieval.
4. Ingredients Shopping	Based on the chosen meals and the ingredients already available at home, the kitchen environment creates a shopping list. The shopping list can be printed or sent to a mobile phone.
5. Nutrition Monitoring	The persuasive kitchen environment tracks a user's diet, evaluates consumption habits, and informs the user about nutritional needs. A mobile application supports the user in tracking his consumption outside his home.

II.2 Related Work

Strictly speaking, we investigate user acceptance towards a persuasive kitchen environment that consists of the five functional scenarios outlined in Table 1. In this section we review the literature on the theoretical foundations of our research as well as academic and industrial research activities that relate to applications in the home appliances domain. We first summarize major literature streams on technology acceptance research and then turn to research work that relates to our persuasive kitchen scenario.

II.2.1 Theoretical Foundation

Research on user acceptance of information technology originates from different theoretical disciplines such as psychology, sociology, and information systems. Various alternative approaches have been proposed to analyze the acceptance and use of a new technology. The majority of technology acceptance models are based on the *Theory of Reasoned Action* (TRA) (Fishbein and Ajzen 1975) and the *Theory of Planned Behavior* (TPB) (Ajzen 1991) . TRA posits that an individual's intention towards a specific behavior can be considered as a predictor of the behavior itself. Antecedents of behavioral intention in the TPB are the attitude towards the behavior, subjective norms, and perceived behavioral control. The *Technology Acceptance Model* (TAM) (Davis 1989) has become the most prevalent model for studying user acceptance in the field of information technology. TAM includes two major predictors of the dependent variable *Behavioral Intention*, which TRA assumes to be closely linked to actual behavior: *Perceived Ease of Use* and *Perceived Usefulness*. More recently, the Unified Theory of Acceptance and Use of Technology (UTAUT) has been proposed, which integrates TAM with other technology acceptance research streams (Venkatesh et al. 2003). UTAUT represents a parsimonious but still comprehensive framework to provide an understanding of factors that affect technology acceptance, and could be confirmed in a large number of research works (see Sun and Zhang (2006) for a review). UTAUT will be applied as the basis for the research model developed in this study.

II.2.2 Application-specific Research

Regarding empirical acceptance studies, there are only a relatively small number of prior studies that investigate user acceptance of concepts similar to the proposed persuasive kitchen environment. So far, research has mainly focused on nutrition (Hanson-Smith et al. 2006), recipe planning (Ju et al. 2001), or communication (Bauer et al. 2005). Although having tested early prototypes with users, these studies are not based on the analysis of larger samples. The only exception we are aware of is a user acceptance study by Rothensee (2008) concerning a simulated 'smart fridge', which offers various assistance functions (product information, automatic replenishment, recipe planner). The results indicate that *Perceived Usefulness* is the strongest predictor to *Behavioral Intention*, followed by emotional response to the product. A significant role of moderating factors (gender, technological competence, sense of presence in a simulation) could not be confirmed.

Besides these few examples of behaviorist research, a number of more technology-oriented projects are known from literature, which confront users with persuasive environments in a realistic home environment. The *Service Centric Home* (Blumendorf et al. 2008) is a research initiative in the course of which a cooking assistant was developed to provide users with cooking recipes and guide the cooking process with videos and explanations. Although the pilot incorporates some connectivity between different kitchen devices and provides some context awareness features, the scenario stays behind a fully integrated kitchen environment and furthermore lacks of persuasive design objectives. A similar approach with regard to home automation is taken by the *Aware Home Research Initiative* at the Georgia Institute of Technology (Kidd et al. 1999). The project explores Ubiquitous Computing applications in the areas of 'Chronic Care Management in the Home', 'Future Tools for the Home', and 'Digit-al Entertainment and Media'.

Finally, several studies provide valuable insights for the application of technology acceptance models in our domain. Garfield (2005) presents results from a longitudinal, qualitative study of the acceptance of Tablet PCs based on interview data from four industries. Main findings include a list of factors that influence the predictors of *Behavioral Intention* in the UTAUT model as well as the identification of the technology's impact on work processes. Ferneley and Light (2008) investigated the role of 'Bystanders' (i.e., persons who are exposed to a technology but are not intended to use it) and conclude that this group may play an important role in technology proliferation. Sheng et al. (2008) studied interaction effects of personalization and context on the intention to adopt. They conclude that increasing personalization raises privacy concerns, and the degree of this relationship is moderated by situational context.

II.3 Research Design

Based on the 'Unified Theory of Acceptance and Use of Technology' (UTAUT) proposed by Venkatesh et al. (2003), we developed and empirically tested a structural model for the explanation and prediction of the users' intention to use a persuasive kitchen environment.

In this section, we describe the research model underlying the study as depicted in Figure 2. Our research objective is to analyze the user acceptance of a persuasive kitchen environment as an example of a persuasive environment in the domestic domain. The most obvious choice regarding the theoretical framework for a study like ours seems to be the classical TAM, which has been used as the foundation for several IS acceptance studies in recent years. For the present study however, TAM may have only limited ability to explain the acceptance of persuasive environments because it neglects the social context in which a technology is being adopted. We consider the social context to be highly important as a persuasive kitchen environment may require behavioral adaptations by all people in the household. For this reason, we decided to construct and test a research model on the foundation of the more comprehensive UTAUT framework and its constructs as proposed by Venkatesh et al. (2003).

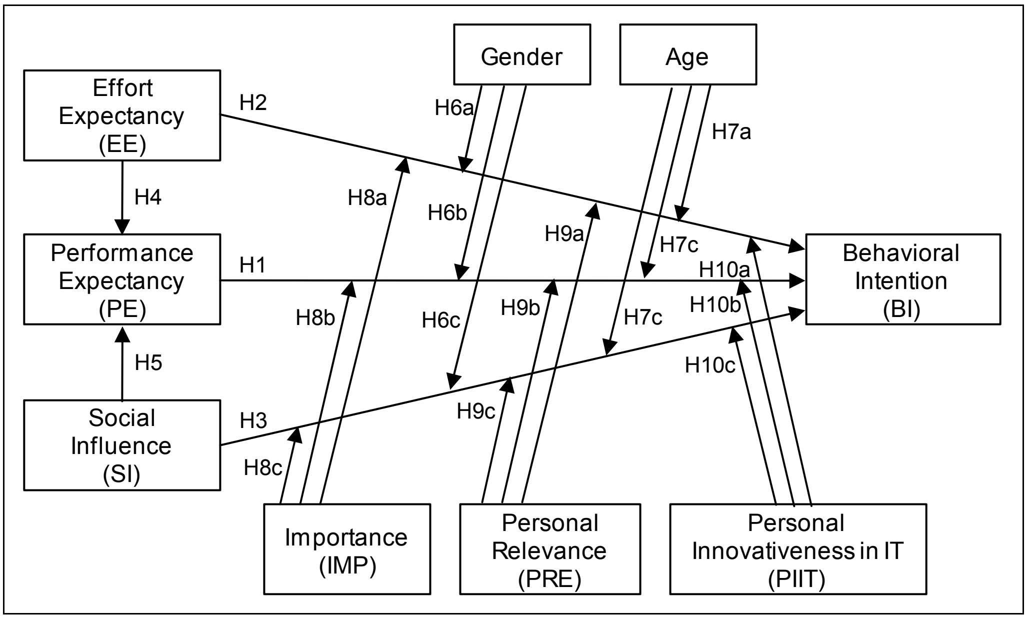


Figure 2: Research Model

Whereas UTAUT has served as the theoretical foundation to many analyses - particularly in industrial settings - it has not yet been applied specifically to persuasive environments in the domestic domain. Furthermore, moderator variables proposed in the original model are not specifically targeted to the typically voluntary use of the investigated application, which is usually the case in the private domain. While basic technology acceptance models have largely matured, the investigation of moderator effects to understand external factors that influence adoption decisions is still underdeveloped and needs to be further elaborated (Dabholkar and Bagozzi 2002; Hong et al. 2002; Sun and Zhang 2006). Hence we introduce additional moderating variables to capture consumer traits and external factors that may influence adoption decisions.

The original UTAUT model posits that four independent variables determine an individual's intention to use a technology: *Performance Expectancy*, *Effort Expectancy*, *Social Influence*, and *Facilitating Conditions* (Venkatesh et al. 2003). *Performance Expectancy* is defined as the degree to which an individual believes that using a particular technology will help him to attain performance gains. *Effort Expectancy* is defined as the degree of ease associated with the use of a particular technology. *Social Influence* is defined as the degree to which an individual perceives that important others believe he should use the new technology. *Facilitating Conditions* are defined as the degree to which an individual believes that an organizational and technical infrastructure exists to support the use of the new technology. *Gender*, *Age*, *Experience*, and *Voluntariness of Use* moderate the key relationships in this model. Similar to TAM, UTAUT has mainly been used in professional settings, but it has also been proven to be useful in non-professional domains (Carlsson et al. 2006; Koivumäki et al. 2006).

To adjust the UTAUT model to our research setting, we made the following modifications to the original model:

1. We eliminated the construct *Use Behavior.* Due to the lack of a working prototype, actual *Use Behavio*r of the proposed scenario cannot be observed. However, *Behavioral Intention* has been shown to be a good predictor of actual behavior as posited by the *Theory of Reasoned Action,* which was confirmed in various studies (Yousafzai et al. 2007).
2. We also eliminated *Facilitating Conditions*. It captures the influence of a supportive infrastructure and is not applicable in the present scenario.
3. We added indirect relationships from *Effort Expectancy* and *Social Influence* on *Performance Expectancy* because this relationship was empirically supported by the results from many prior technology acceptance studies (King and He 2006; Lee et al. 2003; Schepers and Wetzels 2007; Venkatesh and Davis 2000).
4. We eliminated two moderators from the original UTAUT model: *Voluntariness of Use* was eliminated, because the adoption of the proposed kitchen environment will, in contrast to workplace settings, usually occur on a voluntary basis. *Experience* was eliminated, because in the original UTAUT study, *Experience* was examined using a cross-sectional analysis from the time of the artifact's introduction to later stages of larger experience. Due to the early stage of development and the unavailability of a prototype, asking respondents at different stages of experience was not feasible.
5. We decided to introduce three additional moderating variables (*Importance*, *Personal Relevance,* and *Personal Innovativeness in IT*), which will be defined and explained below.

With regard to the direct and indirect relationships between the independent and the dependent variable, we hypothesize in accordance with UTAUT:

H1: *Performance Expectancy* has a positive effect on *Behavioral Intention*.

H2: *Effort Expectancy* has a positive effect on *Behavioral Intention*.

H3: *Social Influence* has a positive effect on *Behavioral Intention*.

H4: *Effort Expectancy* has a positive effect on *Performance Expectancy*.

H5: *Social Influence* has a positive effect on *Performance Expectancy*.

Prior studies observed a high variability in the corresponding correlations, which suggests that moderator variables may exert a significant influence (Lee et al. 2003; Schepers and Wetzels 2007; Sun and Zhang 2006). Moderation occurs when the relationship between two variables depends on a third variable such as gender or age. As a consequence, the introduction of moderating factors can improve the often limited explanatory power and inconsistencies in existing technology acceptance studies. We introduce five moderating variables, which we regard as important in the proposed application setting.

Firstly, we consider the differences in acceptance behavior between men and women (Meyers-Levy and Maheswaran 1991; Venkatesh et al. 2000; Venkatesh and Morris 2000). Men have shown to be usually more pragmatic and task-oriented than women. Moreover, men usually feel more comfortable using new technologies. On the other hand, women compared to men have been found to have a higher awareness of other's feelings, and, in turn, are more influenced by others. Therefore, it seems likely that men are more driven by *Performance Expectancy*, whereas women are more driven by *Effort Expectancy* and *Social Influence*. In accordance with the original UTAUT model, we therefore hypothesize that *Gender* plays a moderating role in our research model.

> **H6a:** For women the effect of *Effort Expectancy* on *Behavioral Intention* is higher than for men.
>
> **H6b:** For men the effect of *Performance Expectancy* is higher than for women.
>
> **H6c:** For women the effect of *Social Influence* on *Behavioral Intention* is higher than for men.

As we are investigating a domestic setting, we theorize that it is possible that gender differences differ from findings of previous studies. Men may be more driven by the opinion of their families than women. Therefore we hypothesize that the effect of *Social Influence* on *Behavioral Intention* is higher for men than for women.

> **H6c':** For men the effect of *Social Influence* on *Behavioral Intention* is higher than for women.

Compared with *Gender*, *Age* has received less attention in the existing literature. Young users have been found to be more driven by *Performance Expectancy*, while older users are more driven by *Effort Expectancy* (Morris and Venkatesh 2000; Venkatesh et al. 2003). It has also been found that older users are more influenced by social factors, because affiliation increases with age and older people are more likely to conform to others' opinions (Sun and Zhang 2006). Following the original UTAUT model, we hypothesize:

H7a: For older people the effect of *Effort Expectancy* is higher than for younger people.

H7b: For younger people the effect of *Performance Expectancy* is higher than for older people.

H7c: For older people the effect of *Social Influence* is higher than for younger people.

Beyond the logic of the original UTAUT model, we introduce additional hypotheses regarding the moderating influences of *Importance*, *Personal Relevance*, and *Personal Innovativeness in IT*. We base the constructs *Importance* and *Personal Relevance* on two aspects included in the *Involvement* construct used in earlier studies. Prior work has investigated the role of *Involvement* on consumer decisions (Zaichkowsky 1985). Barki and Hartwick (1994) investigated its role in the context of information systems development. They define *Involvement* as "a subjective psychological state, reflecting the importance and personal relevance of an object or event". We argue that, following this definition and subsequent applications of the construct, involvement encompasses two different but important factors that influence technology adoption, namely *Importance* and *Personal Relevance*. In the context of our kitchen scenario, *Importance* denotes the extent of intrinsic desire or personal need for support throughout the preparation of a meal. In contrast to that, *Personal Relevance* denotes an individual's general dedication and interest in the application domain. The construct reflects to which extent cooking in general is relevant to an individual. As such it clearly differentiates from the *Importance* construct. For example, cooking can be very relevant for a person when he is often preparing food. At the same time, getting help in the kitchen may not be important for the same person because he is already very skilled. We therefore propose to split the originally proposed involvement construct into the two aspects *Importance* and *Personal Relevance* by introducing separate constructs.

One of the objectives behind the concept of a persuasive kitchen environment is to help users to select and prepare healthier and more tasteful dishes. We theorize that the more a potential user feels that it is important for him to get support in the kitchen, the more important becomes *Performance Expectancy* as a predictor, whereas the influence of *Effort Expectancy* and *Social Influence* will diminish.

H8a: The effect of *Effort Expectancy* decreases with higher *Importance*.

H8b: The effect of *Performance Expectancy* increases with higher *Importance*.

H8c: The effect of *Social Influence* decreases with higher *Importance*.

We furthermore theorize that higher *Personal Relevance* increases the strength of the effect that *Performance Expectancy* exerts on *Behavioral Intention* because functional aspects will be more important than usability or social aspects. Consequently, the significance of *Effort Expectancy* and *Social Influence* should diminish.

> **H9a:** The effect of *Effort Expectancy* decreases with higher *Personal Relevance*.
>
> **H9b:** The effect of *Performance Expectancy* increases with higher *Personal Relevance*.
>
> **H9c:** The effect of *Social Influence* decreases with higher *Personal Relevance*.

Finally, we add the construct *Personal Innovativeness in Information Technology* (*PIIT*) as a moderating factor to our model. Agarwal and Prasad (1998) introduced this construct as a moderating variable into technology acceptance research. In the context of a novel technology, with which only few people are familiar, it could be expected that innovativeness plays an important role in an individual's acceptance behavior. We therefore theorize that people with different levels of *Personal Innovativeness* show different adoption behavior.

> **H10a:** The effect of *Effort Expectancy* decreases with higher *PIIT*.
>
> **H10b:** The effect of *Performance Expectancy* increases with higher *PIIT*.
>
> **H10c:** The effect of *Social Influence* decreases with higher *PIIT*.

II.4 Data Collection

In this section, we describe the data collection process. First, we turn to the instrument development for operationalizing the proposed research model. Then we describe sample selection and provide demographic information about our sample.

II.4.1 Instrument Development

To test the research model and the associated hypotheses, we designed a questionnaire on the basis of existing scales from the technology acceptance literature. The measurement scales for the main constructs were operationalized by adopting items from Venkatesh et al. (2003) and adapting them to the specific context of our persuasive kitchen environment. For constructing measurement scales for *Importance* and *Personal Relevance*, we referred to Barki and Hartwick (1994) and Zaichkowsky (1985).

Personal Innovativeness in Information Technology was operationalized by adapting the scales from Agarwal and Prasad (1998).

Table 2: Measurement Instruments

Construct	Item	Question
Effort expectancy (EE)	EE1	The product concept appears easy to learn.
	EE2	The product concept appears easy to use.
	EE3	I think that I would always feel in control while using this product concept.
Performance expectancy (PE)	PE1	The product concept presented is attractive.
	PE2	The product concept will make preparing food more convenient.
	PE3	The product concept will make food preparation more fun.
Social influence (SI)	SI1	My friends and/or family would be impressed to hear that I use this product concept.
	SI2	My friends and/or family would be interested to use this product concept.
	SI3	I think that using this product concept would enhance my image.
Behavioral intention (BI)	BI1	I would like to have the product concept in my kitchen.
	BI2	I would use the product concept frequently.
	BI3	I wish that the product concept was already available.
Importance (IMP)	IMP1	I often have no idea what to make from the ingredients I currently have in my fridge.
	IMP2	I often end up making something that I have made many times before, because I don't have any inspiration for something new.
	IMP3	I would like to have some assistance when cooking new or difficult dishes.
Personal Relevance (PRE)	PRE1	I really enjoy cooking.
	PRE2	Cooking at home is important to me.
	PRE3	When I am cooking, I feel creative.
Personal Innovativeness in IT (PIIT)	PIIT1	I feel comfortable using new technologies.
	PIIT2	I like to experience products that use new technologies.

As the described persuasive kitchen environment is not yet physically available, we have taken a scenario-based approach. For each of the five functional blocks, we developed a narrative textual scenario description, which was complemented by a graphical illustration created by a professional graphics designer (Table 3).

Table 3: Scenario Descriptions

Scenario	**Description**
Scenario 1: Meal Recom-mendation	 Anna has decided to prepare an evening meal for her family. Although she hopes that the meal she makes is both nutritious and tasty, she doesn't really know which dish she should make. Anna turns on her Cooking Guide and uses it for inspiration on what to cook, Anna interacts with the Cooking  Guide and it proposes several meals, which would be suitable, based upon the ingredients and utensils available in her home and the nutritional needs of her family. Anna browses the options and chooses to prepare fish with lemongrass vegetables.
Scenario 2: Cooking Guidance	 Anna has decided (using the Cooking Guide's recommendation feature) to prepare a fish with lemongrass vegetables dish for her family and some guests that are visiting. She has never made the dish before, and doesn't know how to cook it. Anna wants to cook a tasty and healthy meal for her guests, but she is a little bit worried that she might make a mistake and spoil the meal. The Cooking Guide tells Anna the ingredients she needs to collect, and also tells her the utensils and appliances she will need when preparing the meal. When Anna has collected everything, the Cooking Guide walks Anna through each of the cooking steps. She fills the steamer with the fish and vegetables, and automatically, the Cooking Guide communicates with the steamer and configures it  for the meal she is cooking. Since the Cooking Guide will notify Anna when she has to do the recipe's next step, she can now relax with her guests instead of spending all of her time in the kitchen monitoring the progress of her new meal.

<table>
<tr>
<td>Scenario 3: Recipe Memorization</td>
<td>


Anna is now a more experienced cook and likes to experiment with new ingredients and cooking techniques. She often writes down her newly created recipes on paper after having cooked dinner, but memorizing all of the steps she did and keeping track of all the different recipes is often difficult for her.

Over lunch, Anna had a discussion with a Greek friend who indicated that the best way to cook octopus is to bake it and then grill it. Anna decides that she should try to cook octopus in the way suggested by her friend.

That evening, when Anna begins cooking her evening meal, she indicates to the Cooking Guide that she would like it to memorize the meal she is about to make. The Cooking guide monitors all of Anna's cooking actions and stores this new recipe so that it can be retrieved and used later on.


</td>
</tr>
<tr>
<td>Scenario 4: Ingredients shopping</td>
<td>
Fred likes to cook and he enjoys buying fresh ingredients. Sometimes, when Fred decides to prepare a particular dish and goes shopping to buy what he needs, he returns home and realizes that he has forgotten an important ingredient. He does not like the hassle of having to maintain a shopping list, but he also dislikes forgetting to buy an ingredient that he wants to use in a dish.

For Saturday dinner, Fred decides to prepare lamb chop (Greek style) following a recommendation by the Cooking Guide. The Cooking Guide knows which ingredients are needed for the recipe, which of them he already has enough of and also which ones he needs to buy.

The Cooking Guide presents a shopping list to Fred, which includes the items he needs to buy for the meal he selected, and also some other groceries that he added to his shopping list before-hand. Fred transfers the recipe and the associated shopping list to his mobile phone and goes to his local supermarket. As he collects the items, he updates his shopping list.


</td>
</tr>
</table>

Scenario 5: Nutrition Monitoring	 Peter makes a point of keeping track of his personal nutritional needs, and he makes sure that his diet is always balanced and healthy. He likes to have lunch with friends and enjoys cooking for himself in the evening. However, he often has no clear idea about which nutritional needs he still needs to address at the end of the day when making dinner. Today, Peter buys lunch at his company cafeteria and has a pasta dish with a glass of fresh orange juice. When he pays, the lunch items he bought are automatically noted down and recorded.  When Peter returns home in the evening, he asks the Cooking Guide to make a meal suggestion, taking into account the food that he already ate today. From the list of recommendations, Peter selects a Salad Nicoise, which is low in carbohydrates. Peter is confident that his nutritional needs are always being addressed.

For each scenario, interviewees were asked the same set of questions with minor adaptations to the specific context. All items were measured using a seven-point Likert scale. All constructs were formulated in reflective mode (Chin 1998a). To assure content validity, we followed a two-step process. First, each item was reviewed by three industry experts from a home equipment manufacturer and three academic experts in the area of persuasive systems research. This resulted in a small number of changes to the wording and the overall structure of the questionnaire. The revised questionnaire was then circulated among the same group of experts. It was consistently rated as comprehensive and complete. In a pre-test, we asked ten persons to fill in the questionnaire and provide us with feedback, which led to minor changes for reasons of clarity and comprehensiveness.

II.4.2 Sample

The data for the present study were gathered via an online survey, which was designed following the established principles for survey design by Dillman (2007). The survey was online for two months starting from September 2009. The participation was anonymous, voluntary, and there were no rewards for participation, which can be interpreted to mean that there should be no confounding effects from coercing subjects into participation or due to subjects that are just after some reward. The survey took about 25 minutes to complete.

600 people in different European countries were contacted by email, of which 175 completed the survey. The survey was designed in a way that participants had to answer all questions before they were able to submit the questionnaire. After an initial screening of the data, nine cases were removed from the sample, because of certain patterns that suggested unreliable responses (e.g., the same response category was checked for all questions).

Table 4: Sample Demographics

Variable	Category	Frequency	in %	Mode
Gender	Female	77	46	Male
	Male	89	54	
Age	Under 30	64	39	Under 30
	31-40	49	30	
	41-50	37	22	
	51-60	14	8	
	61 or above	2	1	
Living situation	Living alone	19	11	Other
	Living with parents	2	1	
	Living with partner	30	18	
	Living with partner and	37	22	
	Other	78	47	
Number of children	0	111	67	0
	1	17	10	
	2	23	14	
	3	12	7	
	4	2	1	
	5+	1	1	
Days of cooking per week	0	1	1	5-6
	1-2	30	18	
	3-4	36	22	
	5-6	50	30	
	7	49	30	
Number of days consuming convenience food per week	<1	99	60	1
	1-2	49	30	
	3-4	15	9	
	>5	3	2	
Who does most of the cooking	It is equally divided	37	22	Myself
	My housemate(s)	2	1	
	My partner	13	8	
	Myself	92	55	
	Other	22	13	

The resulting sample comprised 166 subjects corresponding to a final response rate of 28%. Table 4 summarizes demographic information about our sample. Gender distribution is almost balanced with 46% of the respondents being female 54% being male.

39% of respondents were younger than 30 years, 30% were between 31 and 40, 22% between 41 and 50, and 9% older than 51 years, which indicates a bias towards younger people. 40% of the respondents live in a family-like situation (i.e. with partner or with partner and children). A majority of 67% of the respondents does not have any children.

We furthermore asked a number of questions to judge the cooking experience and engagement of our sample. 60% of our respondents are cooking more than 5 times a week, and a majority of 60% is eating convenience food more than once per week. 55% of our respondents cook by themselves, and another 22% share cooking equally with their partner.

To judge the sufficiency of the given sample size, statistical power was determined with the software tool G*Power (Faul et al. 2009). As the maximum number of predictors is 3, a sample size of 119 would be required to detect even small effect sizes above 0.15 with a statistical power of 0.95 at $\alpha=0.05$. With the given sample of n=165, a statistical power of 0.99 is achieved.

II.5 Data Analysis

In this section, we describe the data analysis process and present the results obtained from the study. In the following paragraphs, we describe the applied analysis methodology, the evaluation of measurement and structural model, and the analysis of moderating effects. Furthermore, we compare the five scenarios with each other and point out differences in the evaluation results. We conclude this section with an additional exploratory analysis of certain questionnaire items by means of descriptive statitistics.

II.5.1 Analysis Methodology

II.5.1.1 *Choice of Analysis Method*

The objective of this study is to analyze to which degree users accept the proposed persuasive environment, and to which degree certain factors influence this level of acceptance. For this purpose, we developed the structural model described above, which represents a network of hypothesized relationships between latent variables. Latent variables (or constructs) cannot be observed directly. Therefore the constructs were operationalized by defining questionnaire items. As all constructs have been derived from previous related work, they were operationalized by adopting the original items to our specific context. Next, structural equation modeling (SEM) techniques were applied to test whether the hypothesis system derived from theoretical considerations can be confirmed by empirical data.

Traditionally, multiple regression analysis methods have been applied in empirical studies concerned with causal models. These so-called first generation analysis methods (Fornell and Larcker 1981) are not capable of analyzing item loadings and relationships between constructs simultaneously (Gefen et al. 2000). Instead, two separate analysis steps are required, namely a factor analysis to examine item loadings, and a multiple regression analysis to test the hypothesized paths. Furthermore, the analysis of models with more than one dependent variable requires separate analysis runs for each dependent variable.

SEM analysis methods combine factor and regression analysis into one integrated analysis procedure. Models with multiple dependent variables can be estimated in one analysis run. Furthermore, they provide more rigorous quality criteria for assessing the extent to which a proposed research model is reflected in empirical data.

Two different SEM estimation techniques can be distinguished. Covariance-based approaches estimate the model parameters of a structural equation model such that the resulting covariance matrix for the model reflects the empirical covariance matrix as accurate as possible (Herrmann et al. 2006). The statistical objective of covariance-based approaches is to show that all paths in the research model are plausible against the background of an empirical data set (Gefen et al. 2000), which means that the operationalization of the theory is supported and not disconfirmed by the data (Bollen 1989; Jöreskog and Sörbom 1982). Software applications which implement covariance-based approaches are for example AMOS, EQS, LISREL, and M-Plus.

In contrast to covariance-based approaches, variance-based approaches examine whether regression paths and resulting explained variances (R^2) are significant (Gefen et al. 2000). Partial Least Squares estimation (PLS) is the most prominent method from the class of variance-based approaches, and operates by iteratively performing a number of factor and path analyses until the difference in the average R^2 of the latent variables becomes insignificant (Thompson et al. 1995). Model parameters are estimated with the objective to maximize the ratio of explained variances to overall variances of the dependent variables (latent and observed), which means to minimize their residual variances (Chin 1998a; Lohmöller 1989). Subsequently, jackknifing or bootstrapping procedures are applied so that t-tests based on the obtained pseudo-t-values reveal path significances. Software applications that support PLS are for example LVPLS, PLS-Graph, SmartPLS, and SPAD.

PLS algorithms maximize the prediction of each dependent variable locally and are thus prediction-oriented (Jöreskog and Wold 1982). Variable scores are less accurate than in covariance-based approaches. Path weights are systematically underestimated and item loads are overestimated, which makes PLS estimates more conservative than covariance-based estimates (Chin and Newsted 1999). As a consequence, variance-based approaches are recommended rather for prediction-oriented, exploratory studies, whereas covariance-based approaches are usually applied for confirmatory studies with a strong theoretical foundation. This is corroborated by the fact that for PLS procedures, a recognized global goodness-of-fit index has not yet been established, which is a severe drawback for confirmatory studies (Hulland 1999). The quality of a model can be evaluated only locally, whereas for covariance-based approaches, global fit indices like RMSEA, SRMR, CFI, NNFI, χ^2/df, or AGFI are available (Hair et al. 2006; Hu and Bentler 1998; Marsh et al. 1996). Consequently, PLS lacks of objective measures to assess the degree to which a model fits to an empirical data basis.

Nevertheless, PLS approaches have three important advantages in our context. First they require considerably smaller sample sizes. Whereas for covariance-based approaches, a minimum of 200 samples is recommended (Backhaus et al. 2006; Boomsma 1982; Homburg and Klarmann 2006; Scholderer and Balderjahn 2005), for PLS even small samples sizes below 30 can produce reasonable results.

A second advantage is that PLS does not require any specific distribution, whereas covariance-based approaches require multivariate normally distributed data, which is often violated in questionnaire-based survey data (Chin 1998a). Furthermore, PLS is more robust against multicollinearity (Cassel et al. 1999), which turned out to be an issue in our study.

Third, covariance-based approaches require all observations to be independent from each other, which is not required for PLS (Chin 1998a). As we asked our respondents for their opinion about five scenarios and then aggregated them to obtain results for the whole kitchen environment, not all responses are independent from each other. Table 5 summarizes different characteristics of variance- and covariance-based approaches.

Table 5: Characteristics of Variance- and Covariance-based Approaches

Criterion	Variance-based	Covariance-based
Objective	Prediction-oriented	Parameter-oriented
Approach	Variance-based	Covariance-based
Assumptions	Predictor specification (non-parametric)	▪ Multivariate normal distribution ▪ Independent observations
Parameter Estimates	Consistent if sample size and number of indicators are both large	Consistent
Relationship between latent variables and their measures	Formative or reflective	Typically only reflective. Formative only under certain conditions.
Sensitivity to multicollinearity	Moderate	High
Bias tendency	Conservative	Inflationary at low factor loads
Global goodness-of-fit index	Not available	Several indices available
Scale level	No restrictions	At least interval scale
Model complexity	Large	Small to moderate
Calculation effort	Moderate	High
Sample Size	Minimal recommendations range from 30 to 100 cases. However, when moderately non-normal data are considered, a markedly large sample size is needed despite the inclusion of highly reliable indicators in the model.	Recommendations range from 200, 400, 800 to 5000.
Software applications	LVPLS, PLS-Graph, SmartPLS, and SPAD	AMOS, EQS, LISREL, and M-Plus
Sources: (Bliemel et al. 2005; Chin and Newsted 1999; Chin 1998a; Herrmann et al. 2006; Marcoulides et al. 2009)		

Our survey has both confirmatory and predictive elements. However our primary goal was not to test a specific theory, but to investigate how different factors influence user acceptance. Consequently, our study has a primarily predictive character, so that the major advantage of covariance-based approaches - the availability of global goodness-of-fit indices - is less important. At the same time, several assumptions for covariance-based approaches are violated to some degree. Consequently, PLS was applied on our data to assess the proposed research model.

II.5.1.2 ***Analysis Process***

The assessment of the basic structural equation model (without moderator effects) is usually conducted in a two-step process (Chin 1998a; Schloderer et al. 2009). First, the measurement model is assessed by the means of exploratory and confirmatory factor analysis. After validity and reliability of the measurement model are confirmed, the structural model itself is evaluated in a second step by estimating model parameters by the means of PLS. In addition to the model parameters (item loads, path weights, R^2

values, latent variable scores), path significances, effect sizes and predictive relevance are calculated as quality criteria, which are the basis to judge to which degree the proposed model is supported by the empirical data basis.

Then, the hypothesized moderator effects are assessed. For categorical moderator variables (like gender and age group), an adapted t-test is applied to test for significant interaction effects (Ahuja and Thatcher 2005; Keil et al. 2000; Rai and Keil 2008; Venkatesh and Morris 2000). For continuous latent variables (like *Personal Relevance, Importance, PIIT)*, the so-called product-indicator approach for PLS (Chin et al. 2003) is applied.

For evaluating the measurement model, structural model and moderator effects as described above, all five scenarios were aggregated on the item level. To assess whether the results are consistent across the scenarios, we next repeated the evaluation of the model for each scenario and applied a multi-group comparison to test whether there are significant differences.

Finally we conducted an exploratory data analysis to gain further insights with practical relevance. All scenarios were analyzed by the means of descriptive statistics. Furthermore, we conducted a correlation analysis for several hypothesized relationships that might be interesting from a practical perspective. Figure 3 summarizes the analysis process. In the following sections, the PLS analysis steps for assessing the measurement model, structural model, and moderator effects are described in detail.

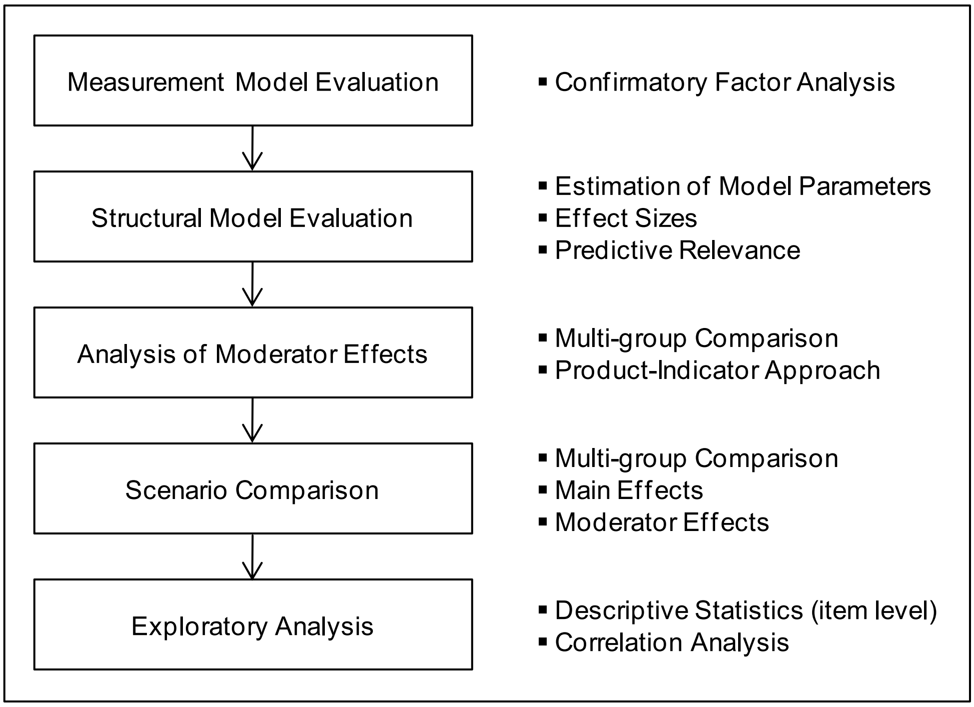


Figure 3: Data Analysis Process

II.5.1.2.1 Measurement Model Evaluation

Constructs in the measurement model are evaluated by assessing different reliability and validity criteria. Reliability "refers to the ability to obtain similar results by measuring an object, trait, or construct with independent but comparable measures" (Churchill 1988). Validity describes the correctness and quality of a measure and is confirmed "when the differences in observed scores reflect true differences in the characteristic one is attempting to measure and nothing else" (Churchill 1979). For reflective measurement models, the following reliability and validity aspects are usually assessed (Hair et al. 2011; Homburg and Giering 1996):

- Indicator reliability
- Convergent validity
- Discriminant validity

Indicator reliability is a measure for the degree to which an indicator variable is an appropriate measure for a latent variable. It indicates which share of the variance of an indicator is explained by its associated latent variable. Practically, indicator reliability is equivalent to the squared standardized factor load. As half of the variance of an indicator should be explained by its associated latent variable, a minimum of 0.707 $(\sqrt{0.5} \approx 0.707)$ is required for each factor load (Hair et al. 2011; Johnson et al. 2006). An often recommended t-test to assess the significance of indicator loads can be omitted under these conditions, as factor loads above 0.707 will always be significant when taking into account the required sample sizes.

Convergent validity measures "the degree to which two or more attempts to measure the same concept through maximally dissimilar methods are in agreement. If two or more measures are true indicators of a concept, then they should necessarily be highly correlated" (Bagozzi and Phillips 1982). It thus indicates to which degree a construct is measured by all its associated factors together (Homburg and Giering 1996). It can be confirmed when the indicators associated with the same construct are highly positively correlated (internal consistency). Construct reliability is measured by *Cronbach's* α and *Composite Reliability* ρ (also named factor reliability, Jöreskog's ρ), and the *Average Variance Extracted* (AVE).

Cronbach's α represents the most commonly used reliability coefficient (Homburg and Giering 1996). There is a lack of consent regarding which threshold a Cronbach's α coefficient should exceed in order to indicate sufficient internal consistency. Nunnally and Bernstein (1994) propose a threshold of 0.7, whereas Malhotra and Birks (2003) propose a threshold of 0.6. Referring to the fact that Cronbach's α increases with the number of indicators associated with a construct, Bagozzi (1980) proposes a limit of 0.5 for two indicators, 0.6 for three and 0.7 for more than three indicators per construct. For this study, we apply a threshold of 0.7, which is most commonly applied in research practice.

Composite Reliability ρ also measures internal consistency, but does not assume that all indicators are equally reliable. This assumption is dropped by taking measurement errors into account. As PLS prioritizes indicators according to their reliability during model estimation, composite reliability is more appropriate for PLS than Cronbach's α (Hair et al. 2011). Values of 0.6 to 0.7 are acceptable in exploratory research, whereas values of 0.7 to 0.9 should be obtained in more advanced stages of research (Nunnally and Bernstein 1994).

Average Variance Extracted (AVE) is a measure for the relation between the explained variance and the measurement error of a latent variable. A threshold of 0.5 is usually recommended to ensure sufficient convergent validity (Chin 1998a; Hair et al. 2011), which means that at least half of the variance of a construct is explained through its associated indicators. Compared to Composite Reliability, AVE is the stronger criterion because ρ can produce a positive result even if more than half of the variance of a construct can be traced back to measurement errors (Chin 1998a).

Discriminant Validity measures the "the degree to which two or more attempts to measure the same concept through maximally dissimilar methods are in agreement. If two or more measures are true indicators of a concept, then they should necessarily be highly correlated" (Bagozzi and Phillips 1982). It is assessed by applying the Fornell/Larcker criterion, which states that the average AVE of a factor must be larger than the squared correlation of this factor with any other factor (Fornell and Larcker 1981; Hair et al. 2011; Homburg and Giering 1996). Whereas the Fornell/Larcker criterion is a very strong measure for discriminant validity (Homburg and Giering 1996), a more relaxed criterion is to assess whether all indicators' loadings with their associated construct are higher than their cross-loadings with any other construct (Hair et al. 2011).

Table 6 summarizes the different criteria to assess the validity of a measurement model. All but the Fornell/Larcker criterion are automatically calculated by the applied software SmartPLS. The Fornell/Larcker criterion has to be assessed manually. If all criteria are fulfilled, the measurement model can be regarded as valid, which is a necessary condition for a valid assessment of the structural model. It shall be noted that for measurement models with formative indicators, different assessment criteria have to be applied (formative indicators are not included in the present research model). As formative indicators associated with the same latent variable do not necessarily measure the same content, and are thus not necessarily correlated highly, they can only be assessed according to their outer weights, their multicollinearity, and their error terms (Hair et al. 2011).

Table 6: Evaluation Criteria: Measurement Model

Scope	Criteria	Threshold
Indicator Reliability	Factor loads	≥ 0.707
Convergent Validity	Cronbach's α	≥ 0.7
	Composite Reliability ρ	≥ 0.6 (exploratory research) ≥ 0.7 (advanced research)
	Average Variance Extracted (AVE)	≥ 0.5
Discriminant Validity	Fornell/Larcker	The average AVE of a factor must be larger than the squared correlation of this factor with any other factor
	Cross loadings	The correlation of each indicator with its associated construct must be larger than its correlation with any other construct.

II.5.1.2.2 Structural Model Evaluation

Once a valid estimation of the constructs is confirmed, the structural model can be assessed according to certain evaluation criteria. R^2 values and the level and significance of the path coefficients are the primary evaluation criteria (Chin 1998a). Additional criteria are effect sizes of the exogenous variables and their predictive relevance.

The *coefficient of determination* R^2 represents the proportion of the total variance of an endogenous variable that is explained by its related latent variables (Backhaus et al. 2011). Hair et al. (2011) note that different research disciplines tend to regard different levels of R^2 as sufficient. Whereas in consumer behavior research, values above 0.2 are regarded as high, in marketing research values of 0.75, 0.50, and 0.25 are usually considered as "substantial", "moderate" and "weak" respectively. In a technology acceptance study, Chin (1998a) considers values of 0.67, 0.33, and 0.19 as "substantial", "moderate" and "weak" respectively. Due to its widespread use, we apply the latter recommendation from Chin (1998a) as guideline for evaluating R^2 values in the present study.

PLS path coefficients can be interpreted similar to the standardized beta coefficients of Ordinary Least Squares (OLS) regression analysis (Krafft et al. 2005).They are assessed with regard to their absolute value, significance and sign. Values close to 1 (or -1) imply a strong influence of a latent variable on their causal successor, whereas values close to 0 indicate weak influence. Values above 0.2 (or below -0.2) can be regarded as substantial (Chin 1998b). Path significances have to be obtained from a bootstrapping procedure because PLS does not assume any specific distribution of the empirical data. Nonparametric bootstrapping allows for conducting a pseudo t-test on the basis of repeated random sampling. Path signs are assessed with regard to their consistency with the underlying hypotheses. Significant paths with a substantial path coefficient and sign, which is directed as expected, are interpreted as empirical support for the underlying hypothesis (Krafft et al. 2005).

The *effect size* f^2 indicates to which degree an exogenous latent variable has substantial impact on an endogenous latent variable. For determining f^2, the structural model is first estimated both with and then without the construct under consideration. Two R^2 values are obtained for each endogenous latent variable: R^2_{included} and R^2_{excluded}. Then the effect size is calculated as:

$$f^2 = \frac{R^2_{included} - R^2_{excluded}}{1 - R^2_{included}}$$

Values above 0.02, 0.15, or 0.35 indicate whether an exogenous latent variable has small, medium or large impact on a related endogenous latent variable (Chin 1998a; Cohen 1988). Effect sizes are not provided automatically by SmartPLS but have to be calculated manually.

Finally the *predictive relevance* of the model (i.e. its ability to predict) is assessed by calculating the so-called Stone-Geisser's Q^2 (Chin 1998a; Geisser 1974; Stone 1974). High Q^2 values indicate that the model is capable of predicting each indicator of an endogenous latent variable (Hair et al. 2011). The predictive relevance is obtained by applying a blindfolding procedure "that omits a part of the data for a particular block of indicators during parameter estimations and then attempts to estimate the omitted part using the estimated parameters. This procedure is repeated until every data point has been omitted and estimated" (Chin 1998a). Different forms of the Q^2 estimation are available. Following the recommendation from Chin (1998a), the "cross-validated redundancy" option (as implemented in SmartPLS) is applied. Q^2 values > 0 indicate sufficient predictive relevance, whereas values < 0 indicate that the applied model is not capable for predicting the empirical data any better than an ordinary mean estimation. Analogous to the effect size f^2, the relative predictive influence q^2 of a variable can be calculated, which is the influence it exerts on the predictive relevance of an endogenous latent variable (Chin 1998a). For determining q^2 for a certain variable, Q^2 is calculated with and without this variable. Then q^2 is obtained as:

$$q^2 = \frac{Q^2_{included} - Q^2_{excluded}}{1 - Q^2_{included}}$$

Table 7 summarizes the assessment criteria for the structural model. It should be noticed that, in contrast to the measurement model, the criteria for the structural model are less strict and leave room for interpretation. They should rather be regarded as guidelines for interpreting a structural model.

Table 7: Evaluation Criteria: Structural Model

Quality criterion	Description	Typical Recommendation	
Coefficient of determination	Proportion of the total variance of an endogenous variable that is explained by its related latent variables.	$R^2 \geq 0.67$ $0.33 \leq R^2 < 0.67$ $0.19 \leq R^2 < 0.33$	substantial moderate weak
Path coefficient	Reflects a hypothesis. Should be assessed with regard to absolute value, significance, and sign.	Path coefficients > 0.2 Sufficient significance (e.g. $p < 0.05$) Sign in accordance with hypothesis	
Effect size	Influence of an exogenous on an endogenous variable	$f^2 \geq 0.35$ $0.15 \leq f^2 < 0.35$ $0.02 \leq f^2 < 0.15$	large medium weak
Predictive relevance	Capability of the model to predict.	Q2 > 0	predictive relevance confirmed
Relative predictive relevance	Influence of a variable on the predictive relevance of another variable.	$q^2 \geq 0.35$ $0.15 \leq q^2 < 0.35$ $0.02 \leq q^2 < 0.15$	large medium weak

II.5.1.2.3 Analysis of Moderator Effects

Besides the direct effects, the evaluation of which has been explained in the previous section, the proposed research model has been extended by a number of moderator effects. Moderator effects "are evoked by variables whose variation influences the strength or the direction of a relationship between an exogenous and an endogenous variable" (Henseler and Fassott 2010). The evaluation of these effects requires different procedures for categorical (e.g. G*ender*) and continuous latent variables (e.g. *Personal Relevance*).

Categorical Variables

For categorical (and particularly dichotomous) variables, a group comparison approach is usually recommended (Ahuja and Thatcher 2005; Henseler and Fassott 2010; Keil et al. 2000; Rai and Keil 2008; Venkatesh and Morris 2000). The rationale of this procedure is to split the data set into two subsets along the parameter values of the dichotomous variable, and to estimate model parameters for each subset separately. A statistical t-test reveals whether the differences between path weights are significant, which is interpreted as confirmation for a hypothesized moderator effect. Keil et al. (2000) suggest the application of a parametric test, which uses standard errors obtained from bootstrapping as input to the test statistics. The test statistics are calculated as follows:

$$t = \frac{Path_{Sample1} - Path_{Sample2}}{\sqrt{\frac{(m-1)^2}{m+n-2} S.E.^2_{Sample1} + \frac{(n-1)^2}{(m+n-2)} S.E.^2_{Sample2}} * \sqrt{\frac{1}{m} + \frac{1}{n}}}$$

In this formula, *m* and *n* denote the sample sizes of the two groups; $Path_{Sample1}$ and $Path_{Sample2}$ are the path coefficients for the path that is being compared; $S.E.^2_{Sample1}$ and $S.E.^2_{Sample2}$ are the variances in each group for the paths that are compared (obtained from the bootstrapping procedure). The t-statistic is then asymptotically t-distributed with m+n-2 degrees of freedom.

It shall be noted that some researchers consider this procedure to be questionable due to the parametric nature of the applied statistical test (Chin and Dibbern 2010; Henseler 2007). They argue that its inherent assumption of normally distributed data contradicts to the generally distribution-free nature of PLS. Therefore, Henseler (2007) and (Chin and Dibbern 2010) developed different assumption-free group comparison approaches for PLS. However, Qreshi and Compeau (2009) found in simulation studies that the results deviate only marginally from the approach presented above, even if data violates the normal distribution assumption. In addition to that, the same authors found that the group comparison approach described above is quite conservative, as group differences are detected only if they are quite strong. Consequently, moderator effects discovered by this procedure are very likely to exist in reality.

Group comparison procedures can be applied directly only for dichotomous variables. Categorical variables that have more than two possible parameter values (like the age classes our example) must be dichotomized first. Though other approaches can be applied, a popular method is a median-split (Henseler and Fassott 2010). Observations for which the parameter value is above the median are assigned to the "high" group. Observations below the median are allocated to the "low" group.

Continuous Variables

In principle, this procedure could be applied to continuous variables as well by dichotomizing them by means of a median-split. This approach however has severe disadvantages. First, a part of the moderator´s variance is neglected in the analysis, and second, observations close to the mean cannot be assigned to a group unambiguously.

Therefore, continuous moderator variables should be assessed by applying the so-called *product-indicator approach*, which usually produces superior results (Henseler and Fassott 2010).

With the product-indicator approach, a new interaction construct is created. The indicators of this new construct are the products of each indicator of the moderating construct with each indicator of the predictor construct (Chin et al. 2003). The PLS algorithm then delivers the path coefficient for the interaction term, and bootstrapping is applied to assess its significance. However, Carte and Russell (2003) make clear that, unlike main effects, moderator effects should not be evaluated on the basis of path coefficients and significance levels because they can be distorted due to measurement errors, different scale levels, and multicollinearity. Instead, a comparison of the R^2 values with and without the moderator construct reveals whether there is a substantial moderator effect.

For deciding whether a change in R^2 is significant, an F-statistic is calculated according to the following formula (Aiken and West 1991; Carte and Russell 2003; Jaccard et al. 1990):

$$F = \frac{(R_2^2 - R_1^2)/(k_2 - k_1)}{(1 - R_2^2)/(N - k_2 - 1)}$$

R_1^2 and R_2^2 are the R^2 values before and after introducing the interaction term; k_1 and k_2 represent the number of predictors before and after introducing the interaction term; N is the sample size. F then follows an F-distribution with $df_1=(k_2-k_1)$ and $df_2=(N-k_2-1)$ degrees of freedom. A standard F-test reveals whether ΔR^2 is significant, which is interpreted as an indication for a significant moderation effect.

II.5.2 Measurement Model

The first assessment step in structural equation modeling is to evaluate the measurement model by means of certain reliability and validity criteria as explained in the previous sections. Table 8 shows the factor analysis results for the assessment of indicator reliability and convergent validity. All item loadings are well above the threshold of 0.707, indicating that over half of the variance is captured by the latent construct (Chin 1998a; Gefen et al. 2000).

The Cronbach's α and composite reliability values used as measures for internal consistency are well above the recommended threshold of 0.7 for each construct (Nunnally and Bernstein 1994). Convergent validity (Cook and Campbell 1979), which refers to the degree to which the items measuring the same construct agree, is examined by considering the average variance extracted (AVE). Table 8 shows that the AVE values are well above the recommended threshold of 0.5 for all constructs (Fornell and Larcker 1981).

Table 8: Validation of the Measurement Model

Construct	Item	Loading	LV Score	SD	α	CR	AVE
Behavioral Intention (BI)	BI1	0.96	4.24	1.77	0.95	0.97	0.91
	BI2	0.95					
	BI3	0.94					
Effort Expectancy (EE)	EE1	0.90	5.01	1.47	0.82	0.89	0.74
	EE2	0.88					
	EE3	0.79					
Social Influence (SI)	SI1	0.78	3.69	1.74	0.82	0.89	0.74
	SI2	0.91					
	SI3	0.88					
Performance Expectancy (PE)	PE1	0.88	4.68	1.68	0.85	0.91	0.77
	PE2	0.88					
	PE3	0.88					
Importance (IMP)	IMP1	0.76	4.21	1.9	0.73	0.84	0.64
	IMP 2	0.88					
	IMP 3	0.76					
Personal Relevance (PRE)	PRE1	0.91	5.01	1.69	0.83	0.89	0.73
	PRE2	0.86					
	PRE3	0.79					
Personal Innovativeness in IT (PIIT)	PIIT1	0.76	5.40	1.69	0.77	0.87	0.77
	PIIT2	0.98					

SD: standard deviation; LV Score: latent variable score; α: Cronbach's α;
CR: Composite Reliability; AVE: average variance extracted

Discriminant validity, which refers to the degree to which measures of distinct concepts differ, was examined by comparing the correlations between the measurement items of distinct constructs with the squared root of the AVE of each construct (Fornell and Larcker 1981). Table 9 shows that the squared root of the AVE for each construct is higher than its correlations with other constructs indicating satisfactory discriminant validity of our measurement model.

Table 9: Discriminant Validity

Construct	BI	EE	SI	PE	IMP	PRE	PIIT
Behavioral Intention (BI)	**0.98**						
Effort Expectancy (EE)	0.57	**0.86**					
Social Influence (SI)	0.68	0.45	**0.86**				
Performance Expectancy (PE)	0.79	0.58	0.64	**0.88**			
Importance (IMP)	0.39	0.23	0.34	0.40	**0.80**		
Personal Relevance (PRE)	0.09	0.14	0.02	0.02	-0.33	**0.85**	
Personal Innovativeness in IT (PIIT)	0.22	0.22	0.27	0.18	0.09	0.40	**0.88**
Note: The diagonal reports the square roots of the average variance extracted (AVE).							

From the evaluation of the measurement model, we can conclude that all reliability and validity criteria are fulfilled. Consequently the measurement model can be taken as the basis for the evaluation of the structural model, which is presented in the following chapter.

II.5.3 Structural Model

With sufficient evidence from reliability and validity measures, the hypothesized paths and the explanatory power of the model can be analyzed. The explanatory power is examined by inspecting the R^2 values (i.e., the explained variance) of the dependent variables. Chin suggests that R^2 values of 0.67, 0.33, and 0.19 in PLS path models should be regarded as substantial, moderate, and weak, respectively (Chin 1998a).

For the basic model without moderators, Figure 4 shows that we obtained R^2 values of 0.69 for *Behavioral Intention* and 0.52 for *Performance Expectancy*, which can be interpreted as substantial and moderate.

Because PLS does not assume a particular distribution, bootstrapping has been applied as a re-sampling technique to determine the statistical significance of the path coefficients. The obtained pseudo- t-values indicate whether the hypothesis that the respective parameter estimates equal zero must be rejected.

The significance tests conducted on the relationships reveal that all paths are significant. The absolute path weights are sufficiently substantial, although *Effort Expectancy* has a comparably weak influence on *Behavioral Intention*. Though some researchers argue that path coefficients larger than 0.2 indicate a substantial relationship (Chin 1998b), we do not regard the low value of 0.2 as a contradiction to the proposed model, as it is on the same level as the values obtained in previous work (Venkatesh et al. 2003).

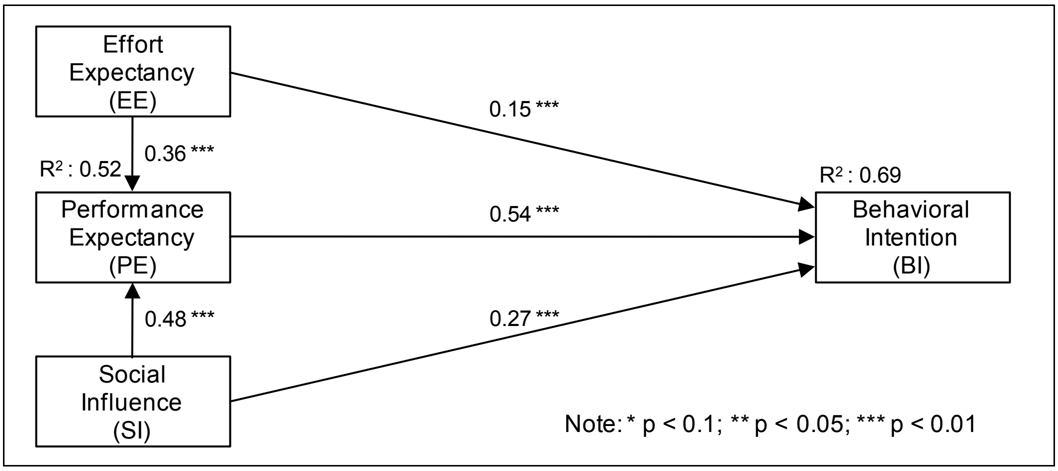


Figure 4: Results for the Structural Model

Three additional criteria shall be applied to evaluate the quality of our model. Effect size f^2 (also named Cohen's f^2) is a measure for the influence of an exogenous on the endogenous variables. Values of 0.02, 0.15, and 0.35 represent "small", "medium", and "large" effect sizes (Chin 1998a; Cohen 1988). Table 10 shows the effect sizes for the individual constructs. *Performance Expectancy* has a large effect on the explained variance of *Behavioral Intention*. *Effort Expectancy* has only a very small effect. *Social Influence* has a small effect, but its effect is much higher than the effect of *Effort Expectancy*. Effect sizes with respect to the *Performance Expectancy* construct are large for *Effort Expectancy* and medium for *Social Influence*. As all values are larger than 0.02, they confirm the appropriateness of the proposed research model.

Table 10: Effect Sizes

Construct	Effect Size f^2 (BI)	Interpretation	Effect Size f^2 (PE)	Interpretation
Effort Expectancy	0.05	small	0.41	large
Performance Expectancy	0.44	large	-	-
Social Influence	0.13	small	0.33	medium

As a second criterion, Stone-Geisser's Q^2 measures the predictive relevance of the model. Values above zero indicate predictive relevance, whereas negative values indicate that the model cannot predict raw data better than a simple mean estimation (Krafft et al. 2005). For the two endogenous variables BI and PE we obtain Q^2 values of 0.62 and 0.38, which are clearly above zero. This indicates that the model has predictive relevance.

By subsequently eliminating each exogenous variable, we obtain a measure for the contribution of each variable to the predictive relevance. In analogy to the calculation of the effect size f^2, we obtain a q^2 value for each exogenous variable. Values of 0.02, 0.15, and 0.35 represent respectively "small", "medium", and "large" predictive relevance (Henseler et al. 2009). The values obtained for our model show that all exogenous variables have at least small predictive relevance (Table 11).

Table 11: Predictive Relevance

Construct	Predictive relevance q^2 (BI)	Interpretation	Predictive Relevance q^2 (PE)	Interpretation
Effort Expectancy	0.06	small	0.32	medium
Performance Expectancy	0.32	medium	-	-
Social Influence	0.14	small	0.20	medium

From our structural model we can conclude that the relations between independent and dependent variables as proposed by our research model are confirmed. Consequently hypotheses H1, H2, H3, H4, and H5 are accepted. Furthermore, we conclude that the model has sufficient explanatory power, which is reflected by substantial effect sizes and predictive relevance.

II.5.4 Moderator Effects

For an examination of moderating effects, we need to distinguish between categorical variables such as *Gender* and continuous latent variables such as *Personal Relevance*, which we measured on a Likert scale. Usually, most researchers have employed an adapted t-test to compare different groups in PLS studies (Ahuja and Thatcher 2005; Keil et al. 2000; Rai and Keil 2008; Venkatesh and Morris 2000). We employ this test to examine the moderation effects of *Gender* and *Age*. The PLS t-test uses the standard errors obtained from bootstrapping in order to test for group equality of path coefficients.

Our finding from this analysis is that only *Gender* exerts a moderating effect on the SI-BI relationship with $p<0.05$ (see Table 12). The direction of the effect is opposed to hypothesis 6c, but in accordance with hypothesis 6c'. All other effects cannot be regarded as significant. Consequently, hypotheses 6a, 6b, 6c, 7a, 7b, and 7c are rejected, and hypothesis 6c' is not rejected.

Table 12: Moderator Effects of Categorical Variables

Moderator		R^2 BI	R^2 PE	Path Coefficients EE → BI	PE → BI	SI → BI
None		0.689	0.517	0.15 ***	0.54 ***	0.27 ***
Gender	Male	0.687	0.556	0.14 ***	0.51 ***	0.32 ***
	Female	0.693	0.480	0.16 ***	0.57 ***	0.21 ***
	T-Test			ns	ns	***
Age	≤ 40 y.	0.686	0.450	0.14 ***	0.54 ***	0.29 ***
	> 40 y.	0.700	0.647	0.17 ***	0.54 ***	0.22 ***
	T-Test			ns	ns	ns
Note: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; ns: not significant						

To test *Importance*, *Personal Relevance*, and *Personal Innovativeness* for their moderating effect, we employed the product-indicator approach (Chin et al. 2003), which is recommended for continuous variables.

Following the results presented in Table 13, we conclude that only *Importance* has a significant moderator effect with regard to the explained variance, although this effect turns out relatively weak if we compare the explained variances with and without the interaction term. Path coefficients of the moderator effect of *Importance* are significant only for two of the three tested relationships at p<0.05, namely the PE-BI and the SI-BI relationship. The effect on EE-BI is not significant.

Table 13: Moderator Effects of Continuous Variables

Moderator		R^2 (BI)	Path Coefficients EE → BI	PE → BI	SI → BI
None		0.6887	0.15 ***	0.54 ***	0.27 ***
IMP	Direct Effect	0.6968	0.14 ***	0.28 ***	0.49 ***
	Interaction	-	0.00	0.45**	-0.38 **
	F-Test	4.30 *	-	-	-
PRE	Direct Effect	0.6919	0.02	0.63 ***	0.25 ***
	Interaction	-	0.19	-0.12	0.02
	F-Test	ns	-	-	-
PIIT	Direct Effect	0.6910	0.22 *	0.59 ***	0.03
	Interaction	-	-0.11	-0.06	0.29*
	F-Test	ns	-	-	-
Note: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; ns: not significant					

Personal Relevance does not significantly improve R^2 and has no significant interaction effects with one of the main relationships. *Personal Innovativeness in IT* has also no significant effect on R^2, which indicates a low effect size. It has a significant influence on the SI-BI relationship, but this influence is opposed to the hypothesized direction. Therefore hypotheses 10a, 10b, and 10c are rejected.

To summarize, hypotheses H8b, H8c are accepted, whereas hypotheses 8a, 9a, 9b, 9c, 10a, 10b, and 10c are rejected.

II.5.5 Scenario Comparison

The presented results are based on aggregated data across all scenarios. Next we investigate to which degree these results are consistent across all scenarios. In the following section the results for the structural model are compared. Then, we explore differences between the scenarios with regard to moderator effects.

II.5.5.1 *Scenario comparison: structural model*

For an assessment of the results for the structural model, we first compare the path weights and R^2 values across the scenarios, and test whether all relationships remain significant. We then apply pair-wise group comparisons to detect statistically significant deviations of path weights between the scenarios.

Table 14 summarizes the path weights, R^2 values and significance levels for all scenarios. Maximum and minimum path weights are highlighted for each relationship. As all paths remain significant across all scenarios, and all R^2 values are reasonably high, we conclude that the proposed structural model with hypotheses H1, H2, H3, H4, H5 can be confirmed for all scenarios.

For the EE-BI relationship, the significance levels are lower for scenario 1 (meal recommendation) and scenario 5 (nutrition monitoring). Regarding the fact that the path weights are also quite low (0.10 and 0.11), we can conclude that *Effort Expectancy* has a comparably weaker influence on the intention to use these scenarios.

Table 14: Scenario Comparison: Path Weights

Parameter	Aggregated	Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5
EE → BI	0.15^{***}	0.10^{*}	0.12^{**}	0.19^{**}	0.18^{***}	0.11^{**}
EE → PE	0.36^{***}	0.26^{***}	0.20^{**}	0.43^{***}	0.42^{***}	0.48^{***}
PE → BI	0.54^{***}	0.56^{***}	0.66^{***}	0.50^{***}	0.47^{***}	0.60^{***}
SI → BI	0.27^{***}	0.26^{***}	0.18^{**}	0.26^{***}	0.32^{***}	0.24^{***}
SI → PE	0.48^{***}	0.51^{***}	0.61^{***}	0.43^{***}	0.42^{***}	0.45^{***}
R^2 (PE)	0.52^{***}	0.43^{***}	0.51^{***}	0.55^{***}	0.53^{***}	0.63^{***}
R^2 (BI)	0.69^{***}	0.64^{***}	0.73^{***}	0.69^{***}	0.70^{***}	0.73^{***}
Note: * p < 0.05; ** p < 0.01; *** p < 0.001						

Furthermore, we can see from Table 14 that path weights and explained variances are different across the scenarios. To judge whether these differences are significant, all parameters were compared across all scenarios, applying the same group comparison approach that was used for the assessment of the categorical moderator variables gender and age (Keil et al. 2000). In Table 15, all paths are listed that differ significantly between two scenarios.

Table 15: Scenario Comparison: Significance of Path Weight Differences

Scenario	Path	2	3	4	5
1	EE → BI				
	EE → PE		*	*	**
	PE → BI				
	SI → BI				
	SI → PE				
2	EE → BI				
	EE → PE		**	**	***
	PE → BI		*	**	
	SI → BI			*	
	SI → PE		**	**	**
3	EE → BI				
	EE → PE				
	PE → BI				
	SI → BI				
	SI → PE				
4	EE → BI				
	EE → PE				
	PE → BI				
	SI → BI				
	SI → PE				
Note: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$					

Whereas between scenarios 3, 4 and 5 path weights are not significantly different, scenarios 1 and 2 both differ in several path weights from scenario 3, 4 and 5. Particularly, the EE-PE relationship seems to be quite unstable, ranging from 0.2 to 0.48. This means that usability aspects determine the expected performance much less for *Meal Recommendation* and *Cooking Guidance* than for *Recipe Memorization*, *Ingredients Shopping*, and *Nutrition Monitoring*.

The relationship PE-BI has a higher path weight for *Cooking Guidance* than for all other scenarios (which is significant in comparison to scenarios 3 and 4). So for the intention to use the different features, the expected usefulness is comparably more important for *Cooking Guidance* than for any other scenario.

The role of *Social Influence* is ambiguous. The direct influence on BI is significantly different only between scenario 2 and 4. Whereas for *Cooking Guidance*, *Social Influence* has only a week influence, for *Ingredient Shopping* this influence is comparably strong. At the same time, *Social Influence* has a comparably strong influence on *Performance Expectancy* in scenario 2. This can be interpreted that for scenario 2, *Social Influence* has a strong influence on the *Perceived Usefulness*, but people base their adoption decision less on what their social environment may think.

Finally it shall be noted that the applied group comparison approach is known to be quite conservative, which means that only remarkable differences become significant (Qreshi and Compeau 2009). Therefore, the scenario differences seem to be more distinct when contemplating the path weights (Table 14) rather than the significance test (Table 15).

II.5.5.2 ***Scenario comparison: moderator effects***

Having compared the path weights, we assess the robustness of the moderator effects across the scenarios. In Table 16, the relevant information is summarized, and significant interaction effects are highlighted.

For the EE-BI relationship, all tested interaction effects are insignificant across the different scenarios. For PE-BI, we found a significant moderator effect by *Importance* in the aggregated scenario. This effect is confirmed to be significant by scenario 3 and 5, and an equally directed, non-significant effect was found for scenarios 2 and 4. Scenario 1 is contradictive to the other scenarios, as there is a negative interaction, yet insignificant and with a comparably weak path weight.

The moderator effect of *Importance* on SI-BI is significantly supported by scenario 3 and 5. All other scenarios support the negative influence, however not on a significant level.

Comparing the path coefficients of the interaction effects with those from the direct effects, we recognize that they are very high, in one case even exceeding 1. This is a clear sign of multicollinearity (Alin 2010; Diamantopoulos and Winklhofer 2001; Grewal et al. 2004). The assumption of significant multicollinearity is supported by an analysis of the Variance Inflation Factors (VIF) for each item. In scenario 5, those items that form the BI construct show VIFs between 8, 9, and 12. There is no consensus among researchers regarding above which VIF value the analysis results become questionable. Henseler et al. (2009) regard values above 10 as critical whereas Diamantopoulos et al. (2008) propose a threshold of 5. So values close or above 10 indicate severe multicollinearity, which may explain the extraordinarily high path weights. The consequence is that the model itself remains valid, but the path weights for the interaction effects should be regarded as critical if the model is used for predictive purposes (Alin 2010).

Table 16: Scenario Comparison: Moderator Effects

Scenario	Moderator	EE → BI	PE → BI	SI → BI	F-Test
Aggregated	Gender	ns	ns	**	
	Age	ns	ns	ns	
	IMP	0.00	0.45**	-0.38**	*
	PRE	0.19	-0.12	0.02	
	PIIT	-0.11	-0.06	0.29*	
Scenario 1	Gender	ns	ns	ns	
	Age	ns	ns	ns	
	IMP	0.26	-0.17	-0.29	*
	PRE	0.34	-0.55	0.38	**
	PIIT	-0.22	-0.42	0.71	*
Scenario 2	Gender	ns	ns	ns	
	Age	ns	ns	ns	
	IMP	0.25	0.35	-0.21	*
	PRE	0.12	0.32	-0.18	*
	PIIT	-0.05	0.11	0.16	
Scenario 3	Gender	ns	*	ns	
	Age	ns	ns	ns	
	IMP	-0.27	0.93**	-0.71**	**
	PRE	0.17	0.14	0.09	*
	PIIT	0.53	-0.06	-0.02	*
Scenario 4	Gender	ns	ns	ns	
	Age	ns	ns	*	
	IMP	-0.19	0.42	-0.07	***
	PRE	0.41	0.06	-0.18	
	PIIT	-0.15	0.12	0.25	
Scenario 5	Gender	ns	ns	ns	
	Age	ns	ns	ns	
	IMP	-0.33	1.11**	-0.78**	***
	PRE	0.06	-0.17	-0.12	**
	PIIT	-0.22	-0.02	0.28	
Note: * p < 0.05; ** p < 0.01; *** p < 0.001; ns: not significant					

II.5.6 Exploratory Analysis

The structural model presented above allows for evaluating different factors that influence the intention to use the different persuasive kitchen scenarios. In this section, these results are complemented by an exploratory analysis of the collected data, in order to get insights into the general acceptance of the scenarios under investigation. The exploratory analysis will be conducted by means of descriptive statistics and graphical illustrations. Originally, all items were measured on a 7-point Likert scale. For the purpose of a concise interpretation, we consider scale values lower than the neutral mean of four as disagreement, and values higher than four as agreement with the respective question.

II.5.6.1 *Attitude towards cooking*

In the first set of questions we asked for the respondent's attitude towards cooking, in order to judge the general relevance of the proposed kitchen environment for the sample population (Figure 5).

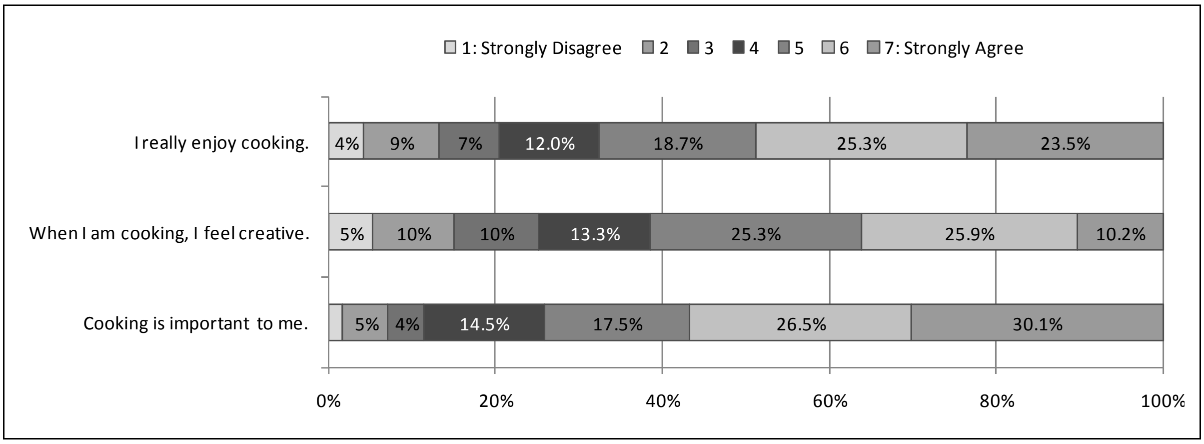


Figure 5: Relevance of Cooking

A majority agrees (with at least 5 points on the Likert-scale) that they enjoy cooking, that they feel creative when they are cooking, and that cooking is generally important to them. Furthermore, we conducted 2-tailed t-tests to check whether the item means significantly deviate from the neutral value of 4 on the Likert scale. Table 17 shows that all item means differ significantly from the neutral value. Since there seems to be a positive attitude towards cooking, we conclude that the proposed scenarios are of sufficient relevance for our sample population.

Table 17: Deviation from Scale Mean

Item	Mean	Deviation from scale mean	Sig. (2-tailed)
I really enjoy cooking.	5.02	1.018	.000*
When I am cooking, I feel creative.	4.62	.620	.000*
Cooking is important to me.	5.40	1.404	.000*

II.5.6.2 *Importance of getting assistance*

The following set of questions was aimed at investigating whether the core functions of the proposed kitchen environment are perceived as being important. The core functions are to provide shopping support, to propose recipes, to guide the meal preparation process, and to help users in adopting a healthier nutrition behavior. Figure 6 shows the results for the respective survey questions, and in Table 18, we apply t-tests to analyze whether these results differ significantly from the neutral scale mean of 4.

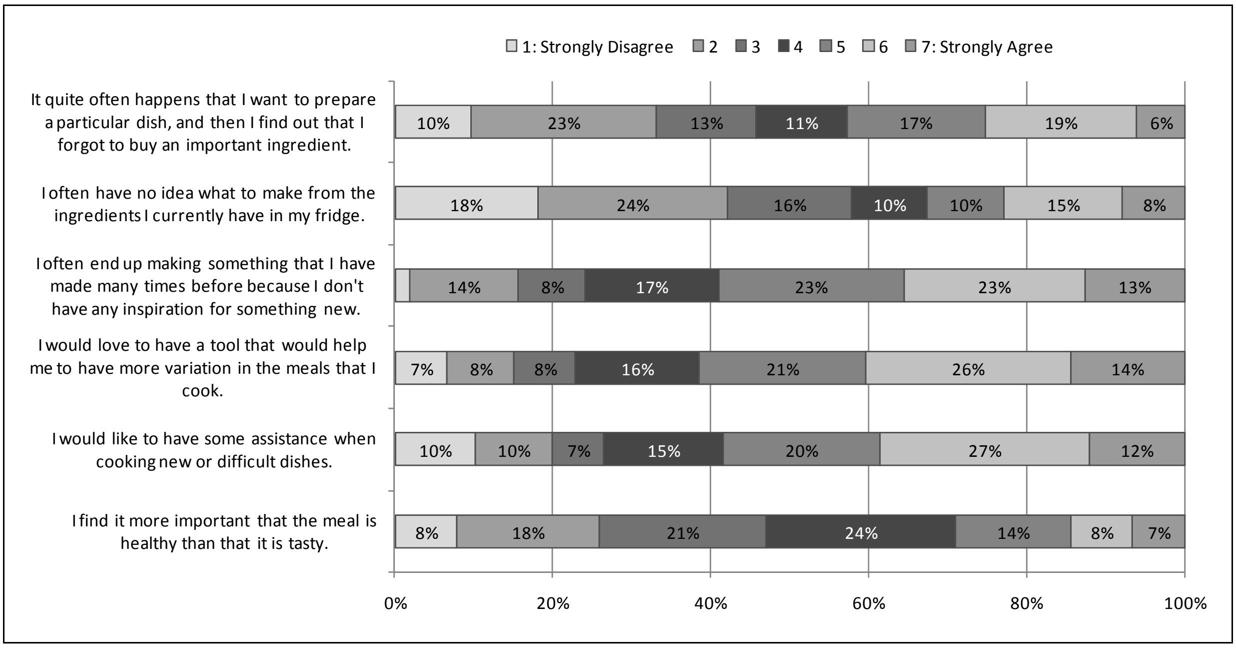


Figure 6: Importance of Getting Assistance

Regarding shopping assistance, the answers are quite balanced across the scale, and the mean of this item does not significantly differ from the neutral scale mean. 42% agree that they often forget to buy an important ingredient (5 or more scale points), and 46% disagree (3 or less scale points). Compared to other core functions, shopping support seems to be of less importance for our sample population, which is further corroborated by the fact that the mean value for this question does not significantly differ from the neutral value. This can be interpreted that respondents are, on average, indifferent about this function.

With regard to the recipe recommendation function, we conclude that our sample population has a strong wish to get inspirations and proposals for new recipe variations, whereas only few people have a general lack of ideas for preparing meals out of a given set of ingredients. Only 33% of the respondents state they often lack an idea of what to prepare, whereas 61% state that they often have no inspiration for a new dish. 59% would even wish to have tool support to get inspirations for more varied dishes. Although both aspects - providing a general idea of what to prepare and providing recommendations on new variations - are perceived positive, the desire for getting inspirations for new variants seems to be stronger. This is corroborated by the fact that for the first aspect, the mean value lies significantly below the neutral scale value, whereas the second aspect is rated significantly better than the neutral mean.

The desire for assistance in preparing new or difficult meals has been rated relatively positive. 59% of the respondents would appreciate such assistance, whereas only 27% disagree.

Finally we asked whether our sample population rates healthiness or tastiness higher. Only 29% agree that healthiness is more important, whereas 47% disagree with this statement.

Table 18: Deviation from Scale Mean

Item	Mean	Deviation from scale mean	Sig. (2-tailed)
It quite often happens that I want to prepare a particular dish, and then I find out that I forgot to buy an important ingredient.	3.86	-.145	.315
I often have no idea what to make from the ingredients I currently have in my fridge.	3.45	-.548	.000***
I often end up making something that I have made many times before because I don't have any inspiration for something new.	4.66	.657	.000***
I would love to have a tool that would help me to have more variation in the meals that I cook.	4.72	.717	.000***
I would like to have some assistance when cooking new or difficult dishes.	4.52	.524	.000***
I find it more important that the meal is healthy than that it is tasty.	3.69	-.307	.015*
Note: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$			

To summarize our results on the respondents' general attitude towards cooking assistance, we can conclude that cooking is generally important for our sample. Respondents appreciate the idea of getting help in finding and preparing new recipes, whereas shopping support is only weakly appreciated. Getting inspiration for tasty meals is a stronger desire than getting proposals for more healthy dishes. These results indicate that our respondents rate assistance features and the hedonic value of the proposed persuasive kitchen environment higher than its capability of supporting users in changing their nutrition habits.

II.5.6.3 *Perception of scenarios*

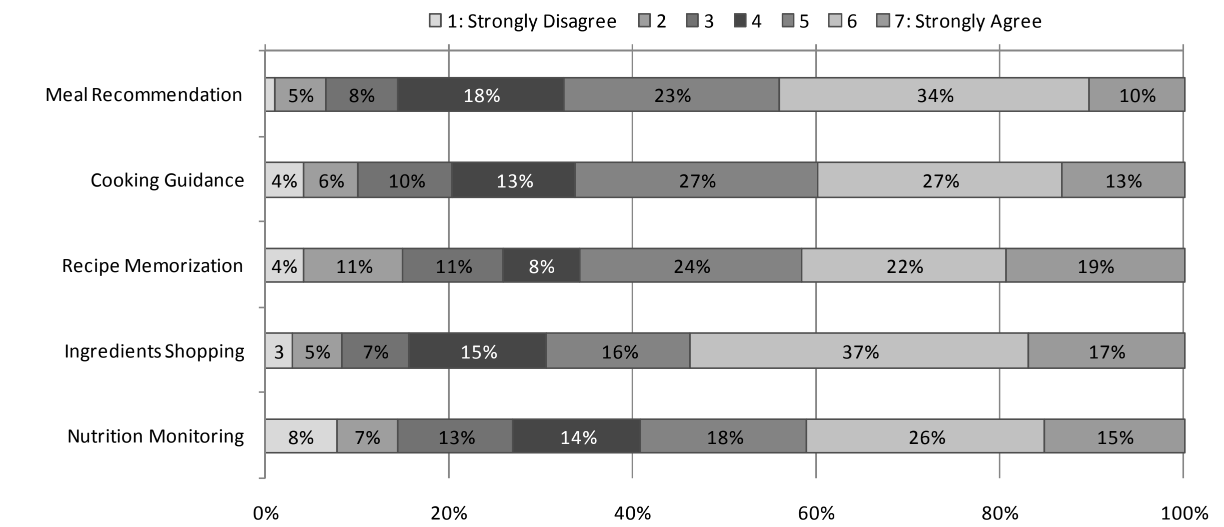
After having investigated the general attitude towards cooking assistance, this section descriptively compares the individual scenarios on the item level. The section is structured along the main constructs *Performance Expectancy*, *Effort Expectancy*, *Social Influence*, and *Behavioral Intention*. For each item, relative frequencies are illustrated for the five scenarios. Furthermore, descriptive statistics are presented, including a t-test-based assessment on whether the responses significantly differ from the scale mean.

II.5.6.3.1 Performance Expectancy

The purpose of the *Performance Expectancy* construct is to assess how respondents judge the usefulness of the proposed scenarios. In the first question, respondents were asked whether they find the presented concepts attractive (Table 19). All five scenarios are perceived attractive, which is reflected in item means well above the neutral value of 4.0 (significant for all scenarios), and modes of 5 or 6 on the 7-point Likert-scale.

Table 19: Perception of Scenarios: PE1

PE1: The product concept presented is attractive.



Scenario	Mode	Mean	Std. Dev.	Deviation from scale mean	Sig. (2-tailed)
Meal Recommendation	6	4.99	1.39	0.99	.000***
Cooking Guidance	5 ; 6	4.84	1.56	0.84	.000***
Recipe Memorization	5	4.81	1.78	0.81	.000***
Ingredients Shopping	6	5.13	1.57	1.13	.000***
Nutrition Monitoring	6	4.66	1.81	0.66	.000***
Note: * p < 0.05; ** p < 0.01; *** p < 0.001					

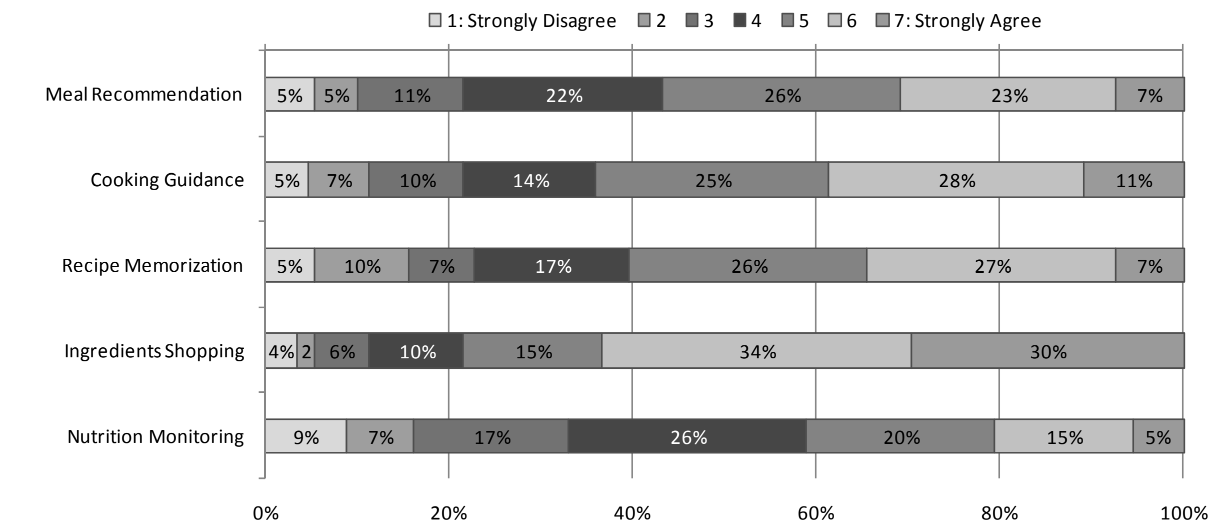
Ingredients Shopping is regarded as most attractive, whereas N*utrition Monitoring* gets the lowest but still positive rating. Between 59% and 68% of the respondents rate the scenarios as attractive (a value of 5 or more on the Likert scale).

Next, respondents were asked whether they expect the proposed concepts to make food preparation more convenient (Table 20). Besides *Nutrition Monitoring,* all scenarios were rated significantly better than neutral. *Cooking Guidance* and *Ingredients Shopping* received the best rating, which could be expected as these functions are aimed at increasing convenience. The comparably lower rating of *Nutrition Monitor-*

ing may be caused by the respondents' expectation that this function requires additional effort for tracking and evaluating one's individual nutrition behavior.

Table 20: Perception of Scenarios: PE2

PE2: The product concept will make preparing food more convenient.



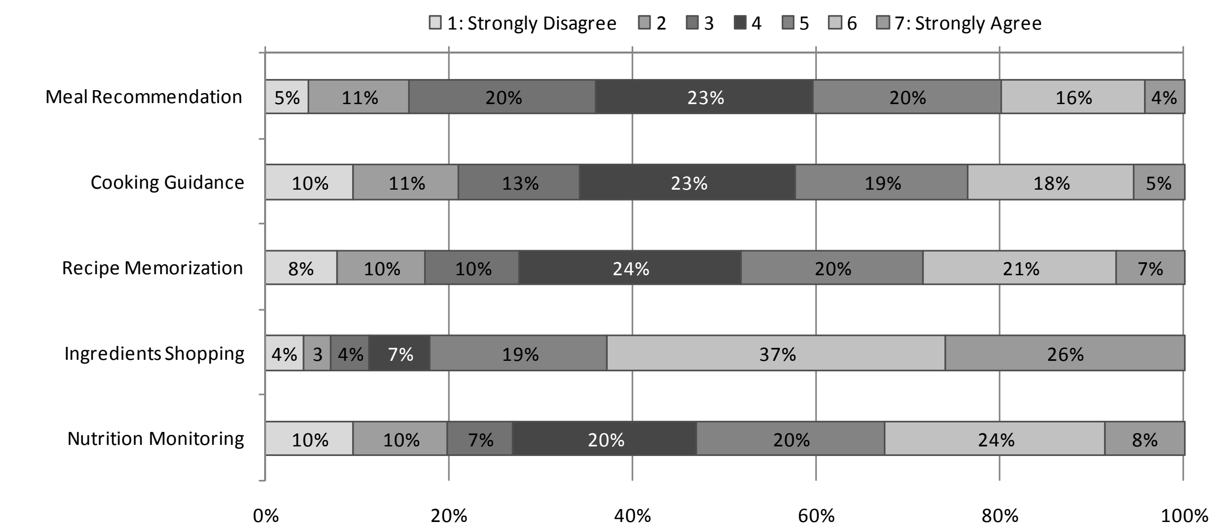
Scenario	Mode	Mean	Std. Dev.	Deviation from scale mean	Sig. (2-tailed)
Meal Recommendation	5	4.57	1.52	.57	.000***
Cooking Guidance	6	4.75	1.61	.75	.000***
Recipe Memorization	6	4.58	1.63	.58	.000***
Ingredients Shopping	6	4.75	1.76	.75	.000***
Nutrition Monitoring	4	4.08	1.61	.08	.499

Note: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

To capture the hedonic dimension of the proposed kitchen environment, we asked whether respondents expect that food preparation will be more fun with the product concept in place (Table 21). Surprisingly, *Ingredients Shopping* received the highest score. 82% of the respondents rated this scenario better than neutral. At the same time, *Cooking Guidance* and *Meal Recommendation* received a comparably low score that does not significantly differ from the neutral mean. The score for *Cooking Guidance* in particular is comparably low; this is surprising because this function was aimed at easing the preparation of new or complicated recipes, which should foster a joyful cooking experience. *Nutrition Monitoring* is regarded as potentially joyful, with a rating slightly above the neutral mean. Evidently, it may be a joyful experience for consumers to gain transparency of their nutrition behavior and get recommendations on how to improve, even though it comes along with some inconveniences.

Table 21: Perception of Scenarios: PE3

PE3: The product concept will make food preparation more fun.



Scenario	Mode	Mean	Std. Dev.	Deviation from scale mean	Sig. (2-tailed)
Meal Recommendation	4	4.08	1.51	.08	.506
Cooking Guidance	4	4.06	1.69	.06	.647
Recipe Memorization	4	4.31	1.68	.31	.019*
Ingredients Shopping	6	5.51	1.54	1.51	.000***
Nutrition Monitoring	6	4.37	1.78	.37	.007**
Note: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$					

To summarize the results for the *Performance Expectancy* construct, the *Ingredients Shopping* scenario consistently received the best rating, which indicates the existence of a significant user need. *Nutrition Monitoring* got relatively low ratings with regard to attractiveness and convenience. *Meal Recommendation, Cooking Guidance and Recipe Memorization* received a positive scoring considering attractiveness and convenience, but are expected to be comparably less joyful. Despite these differences, respondents have relatively positive performance expectancies across all items and scenario.

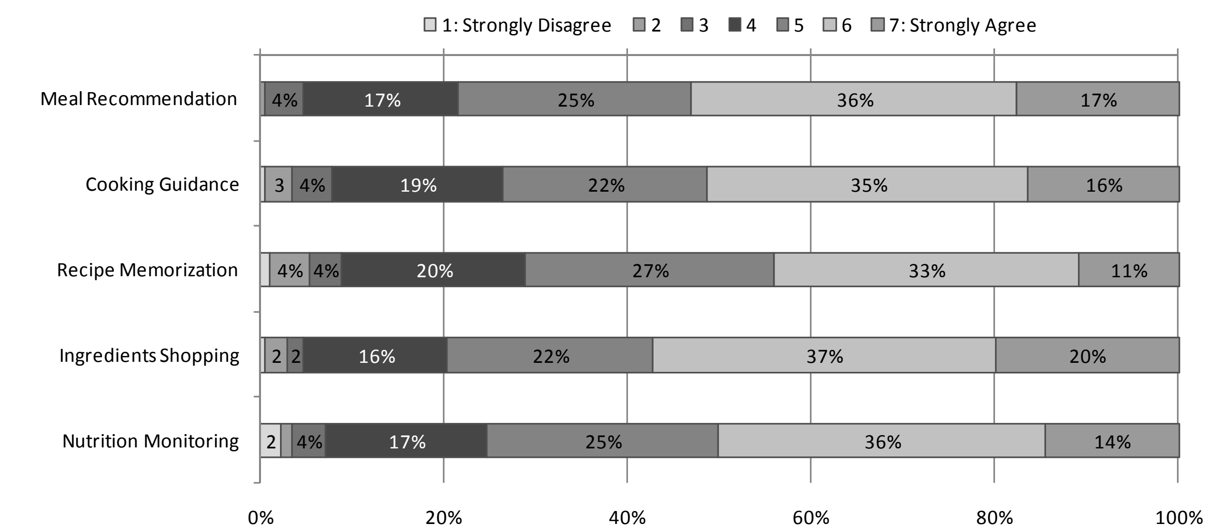
II.5.6.3.2 ***Effort Expectancy***

The *Effort Expectancy* construct is intended to assess how respondents perceive the usability of the proposed scenarios. Three questions were asked with regard to learning effort, ease of use and the feeling of having control over the proposed functionality.

Evidently, respondents do not have major concerns regarding the effort to learn how to operate the persuasive kitchen environment (Table 22). Between 71% and 79% expect that it will be easy to learn how to use the different functionalities. As expected, *Recipe Memorization* has received the lowest score as this will be the most complex individual functionality. *Ingredients Shopping* has received the highest score, which was expected since it will not require users to adapt to new processes, but only to use the recommended shopping list generated by the system.

Table 22: Perception of Scenarios: EE1

EE1: The product concept appears easy to learn.

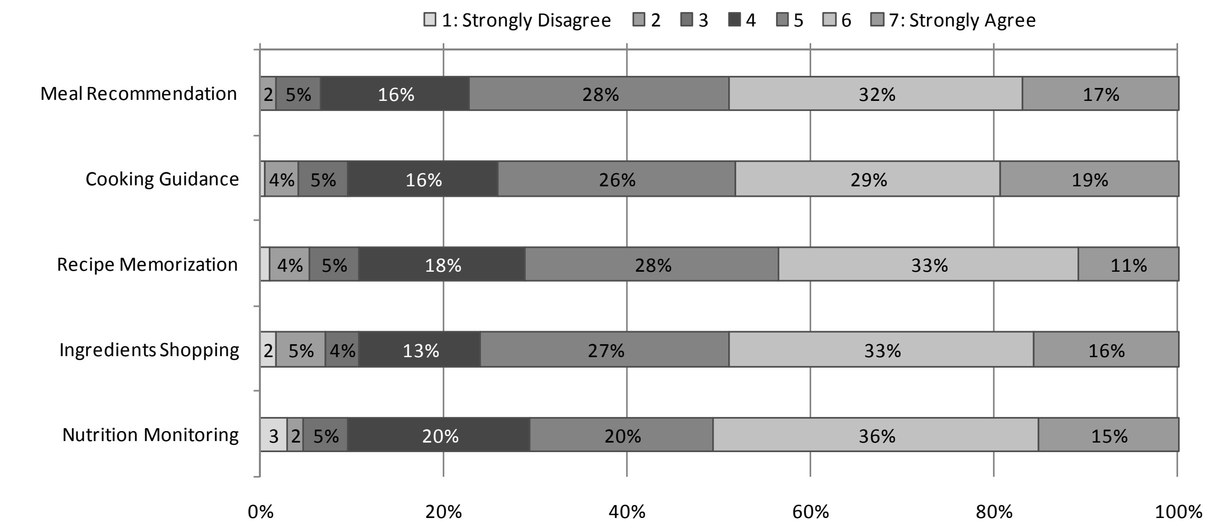


Scenario	Mode	Mean	Std. Dev.	Deviation from scale mean	Sig. (2-tailed)
Meal Recommendation	6	5.43	1.12	1.43	$.000^{***}$
Cooking Guidance	6	5.29	1.28	1.29	$.000^{***}$
Recipe Memorization	6	5.10	1.30	1.10	$.000^{***}$
Ingredients Shopping	6	5.48	1.22	1.48	$.000^{***}$
Nutrition Monitoring	6	5.27	1.30	1.27	$.000^{***}$
Note: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$					

Similarly, ease of use consistently received relatively positive scores for all scenarios (Table 23). Between 72% and 77% of the respondents agreed that the scenarios appear easy to use. Surprisingly, *Recipe Memorization* was rated only slightly weaker than the other scenarios, although this scenario requires one to learn novel operation processes. Either respondents had trust that, although being new and complex, this functionality can be designed in a usable fashion, or they did not recognize the complexity of this feature. *Meal Recommendation* was rated most positive, which could be expected as this scenario appears to be similar to what users know from existing web sites that provide recipe recommendations (in an additional question, respondents were asked whether they are using the internet for recipe searching, which was confirmed by 84%).

Table 23: Perception of Scenarios: EE2

EE2: The product concept appears easy to use.



Scenario	Mode	Mean	Std. Dev.	Deviation from scale mean	Sig. (2-tailed)
Meal Recommendation	6	5.34	1.18	1.34	.000***
Cooking Guidance	6	5.27	1.34	1.27	.000***
Recipe Memorization	6	5.08	1.32	1.08	.000***
Ingredients Shopping	6	5.20	1.41	1.20	.000***
Nutrition Monitoring	6	5.19	1.40	1.19	.000***

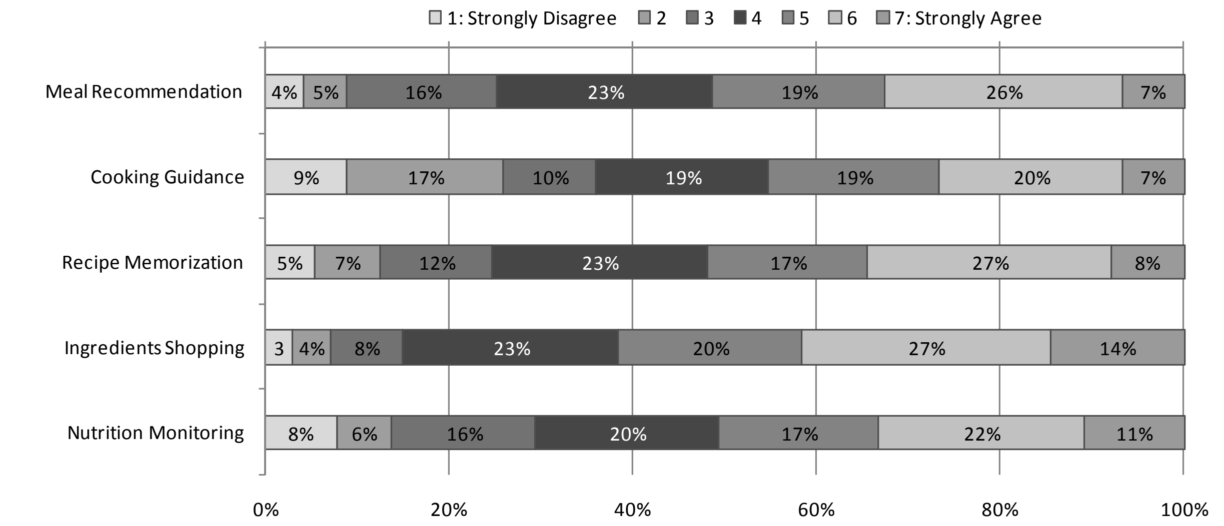
Note: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

In the third question relating to the *Effort Expectancy* construct, we asked whether respondents would feel in control while using the product concept (Table 24). The responses for *Cooking Guidance* do not lie significantly above the neutral level and are the lowest among all scenarios. The logic behind may be that respondents recognize that this functionality interferes with their current habits and takes over part of the control of the cooking process. Considering this fact, it seems surprising that the respondents rated the *Cooking Guidance* scenario at least neutral.

The other scenarios interfere to a lesser degree with current habits. Consequently, they were all rated significantly positive. The highest score was given to *Ingredients Shopping*, which may be regarded as the scenario of lowest interference. *Meal Recommendation*, *Recipe Memorization*, and *Nutrition Monitoring* also achieved significantly positive scores. Comparing the means of this item with the means of *ease of use* and *ease of learning*, the *feeling of control* was rated lower. So respondents may have some concerns whether they can always keep the functionality under control.

Table 24: Perception of Scenarios: EE3

EE3: I think that I would always feel in control while using this product concept.



Scenario	Mode	Mean	Std. Dev.	Deviation from scale mean	Sig. (2-tailed)
Meal Recommendation	6	4.52	1.52	0.52	.000***
Cooking Guidance	6	4.07	1.78	0.07	.602
Recipe Memorization	6	4.51	1.61	0.51	.000***
Ingredients Shopping	6	4.92	1.51	0.92	.000***
Nutrition Monitoring	6	4.43	1.72	0.43	.001***
Note: * p < 0.05; ** p < 0.01; *** p < 0.001					

To summarize the results for the *Effort Expectancy* construct, we can conclude that the respondents do not have major concerns with regard to usability, learning effort, and control of the different functionalities. They seem to have confidence that the scenarios can be implemented such that they integrate more or less seamlessly into their current cooking and shopping habits, and they can imagine to become effective in using the proposed scenarios.

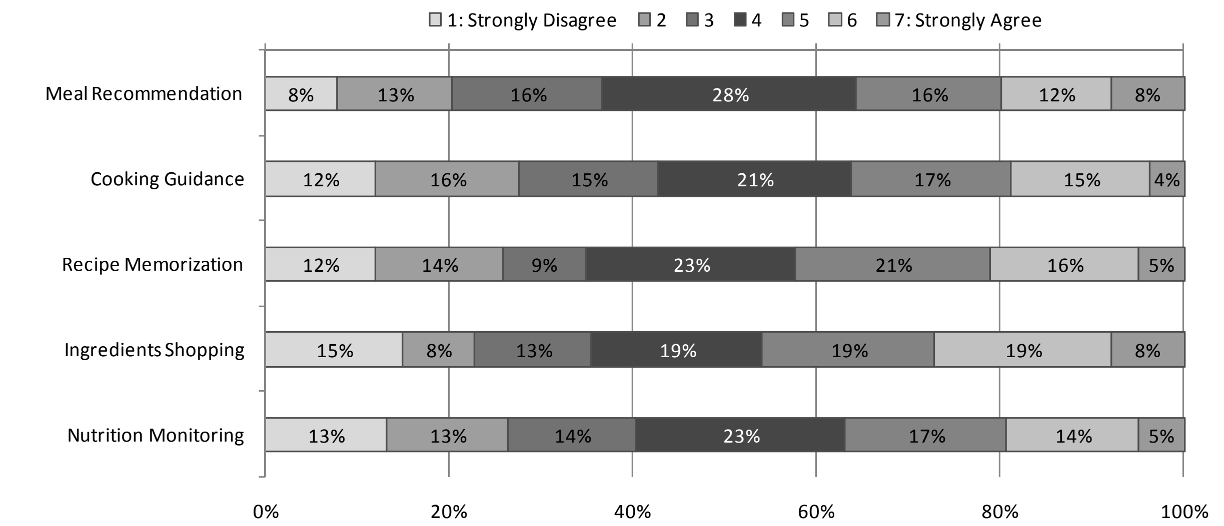
II.5.6.3.3 ***Social Influence***

The *Social Influence* construct is intended to measure how respondents may be influenced by their social environment in their decision to use the proposed concept.

In the first question relating to the *Social Influence* construct, respondents were asked whether they think that their friends and families would be impressed to hear that they use the product concept (Table 25). A significantly positive or negative score was not obtained in any of the scenarios, although there is a slightly negative tendency for four scenarios. The relatively best rating was given for *Ingredients Shopping* (46% positive versus 36% negative answers). *Cooking Guidance* obtained the relatively lowest rating with 36% positive vs. 43% negative responses). It seems that the respondents did not perceive the scenarios as an appropriate means for self-expression towards their friends and families, and they did not consider the proposed kitchen environment as something impressive for their social environment.

Table 25: Perception of Scenarios: SI1

SI1: My friends and/or family would be impressed to hear that I use this product concept.



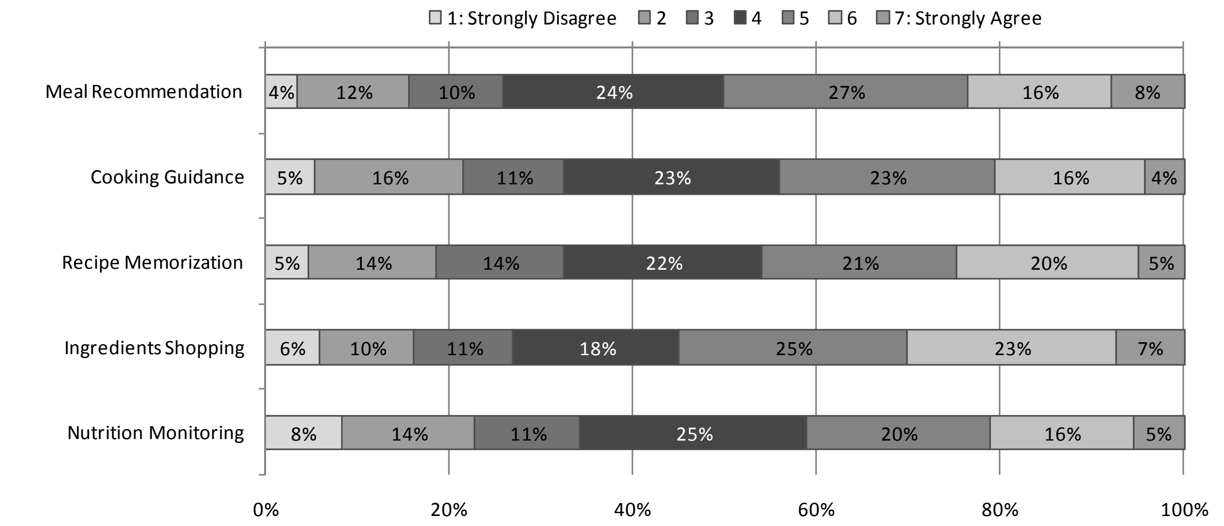
Scenario	Mode	Mean	Std. Dev.	Deviation from scale mean	Sig. (2-tailed)
Meal Recommendation	4	3.98	1.65	-0.02	.888
Cooking Guidance	4	3.76	1.71	-0.24	.071
Recipe Memorization	4	3.95	1.74	-0.05	.722
Ingredients Shopping	6	4.07	1.86	0.07	.618
Nutrition Monitoring	4	3.81	1.74	-0.19	.155

Note: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Being asked whether their friends and families would be interested to use the product concepts, the respondents gave slightly more positive answers (Table 26). *Ingredients Shopping* (55% positive vs. 27% negative responses) and *Meal Recommendation* (51% positive vs. 26% negative responses) were the only two scenarios that were ranked significantly positive. *Cooking Guidance, Recipe Memorization,* and *Nutrition Monitoring* achieved almost neutral scores, which is reflected in the mode value of 4 for these scenarios. In relation to the answers given for other items, the scores for this item are relatively low. At the same time, all scenarios achieved more positive than negative answers, so that we can conclude that the respondents would expect modest interest of their friends and families.

Table 26: Perception of Scenarios: SI2

SI2: My friends and/or family would be interested to use this product concept.



Scenario	Mode	Mean	Std. Dev.	Deviation from scale mean	Sig. (2-tailed)
Meal Recommendation	5	4.36	1.55	0.36	.003**
Cooking Guidance	4 ; 5	4.09	1.59	0.09	.464
Recipe Memorization	4	4.19	1.60	0.19	.121
Ingredients Shopping	5	4.43	1.65	0.43	.001***
Nutrition Monitoring	4	4.02	1.67	0.02	.889

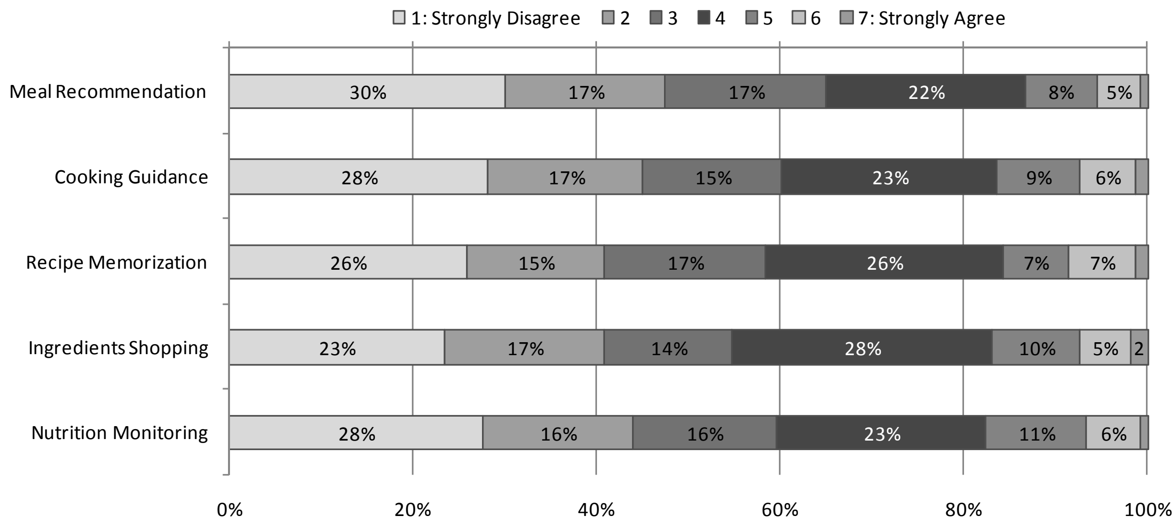
Note: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

The third *Social Influence* item raised the question whether respondents think that the proposed product concept would enhance their image (Table 27). The answers for all scenarios were significantly negative. For four scenarios, the most frequent answer was even "Strongly Disagree". Positive answers remain below 20% for all scenarios, whereas the negative answers range from 54% to 64%. Deviations from the scale mean are relatively high compared to other items, which shows that the respondents were very determined and resolute in their rating.

Interpreting these results neutrally, we might conclude that a persuasive kitchen environment is not regarded as being an appropriate means for enhancing one's social image, which seems to be logical. However, the strong negative tendency may indicate that respondents fear that telling friends about using a persuasive kitchen environment might even have rather negative consequences for their social image. An explanation might be that the scenarios are perceived as over-automation, which may cause concerns that a potential adoption might lead to a kind of "nerd"-image.

Table 27: Perception of Scenarios: SI3

SI3: I think that using this product concept would enhance my image.



Scenario	Mode	Mean	Std. Dev.	Deviation from scale mean	Sig. (2-tailed)
Meal Recommendation	1	2.77	1.55	-1.23	.000***
Cooking Guidance	1	2.91	1.62	-1.09	.000***
Recipe Memorization	1 ; 4	3.00	1.61	-1.00	.000***
Ingredients Shopping	4	3.07	1.60	-0.93	.000***
Nutrition Monitoring	1	2.93	1.61	-1.07	.000***

Note: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Evidently, a persuasive kitchen environment is not perceived as a product that is used for motives such as self-expression or image enhancements. There might even be a tendency to regard the proposed concept as harmful for one's social image (which would require further investigations). Comparing the results for *Social Image* with those for *Performance Expectancy*, we can conclude that whereas usefulness was rated rather positive, social aspects are not a strong argument for our respondents to become interested in the persuasive kitchen scenarios.

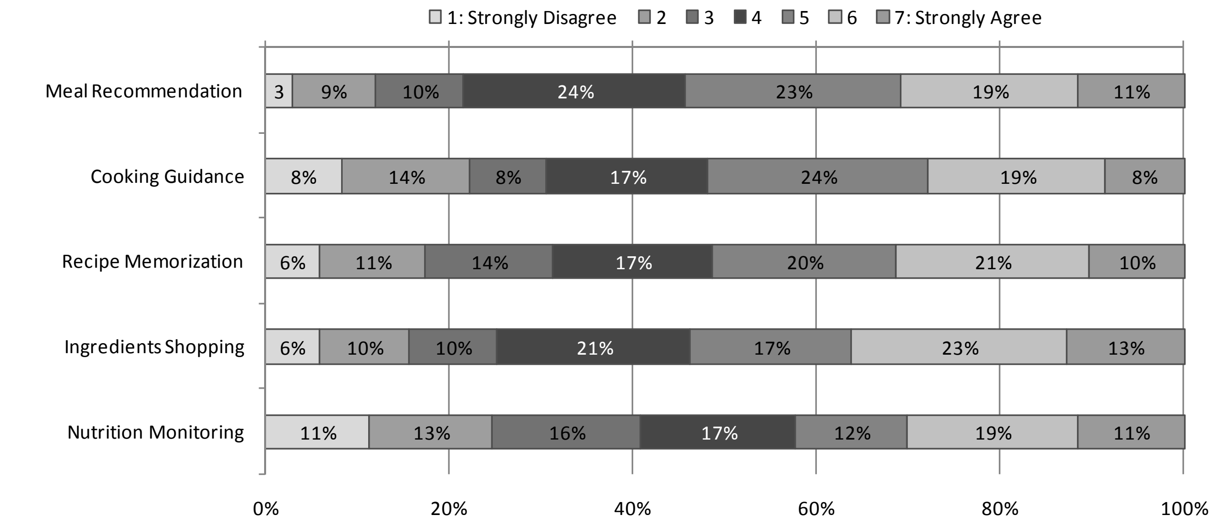
II.5.6.3.4 **Behavioral Intention**

Behavioral Intention measures the intention to use the proposed scenarios. As such, it can be interpreted as a predictor for a final adoption decision, even if several other aspects will influence an ultimate buying decision, which are not considered here. Therefore, the items that form the *Behavioral Intention* construct give indications on how the proposed persuasive kitchen environment might be adopted by consumers once it is available.

In the first question related to *Behavioral Intention*, respondents were asked whether they would like to have the proposed concepts in their kitchen (Table 28). Only three scenarios received a significantly positive scoring. *Meal Recommendation* was ranked best with 43% positive vs. 22% negative answers. *Recipe Memorization* and *Ingredients Shopping* were also perceived significantly positive. *Cooking Guidance* does not differ significantly from the neutral mean but received 41% positive vs. 30% negative answers so that we can still assume a positive tendency. *Nutrition Monitoring* received the lowest rating with 42% positive vs. 40% negative answers. This result was surprising as *Nutrition Monitoring* is the basis for providing support in adopting a healthier nutrition behavior, whereas the other scenarios are aimed at supporting the preparation process. This indicated that supporting healthiness is a weaker argument for a potential buying decision than comfort, tastefulness and pleasure.

Table 28: Perception of Scenarios: BI1

BI1: I would like to have the product concept in my kitchen.



Scenario	Mode	Mean	Std. Dev.	Deviation from scale mean	Sig. (2-tailed)
Meal Recommendation	4	4.60	1.56	0.60	.000***
Cooking Guidance	5	4.27	1.77	0.27	.055
Recipe Memorization	6	4.38	1.73	0.38	.005**
Ingredients Shopping	6	4.55	1.73	0.55	.000***
Nutrition Monitoring	6	4.07	1.91	0.07	.655

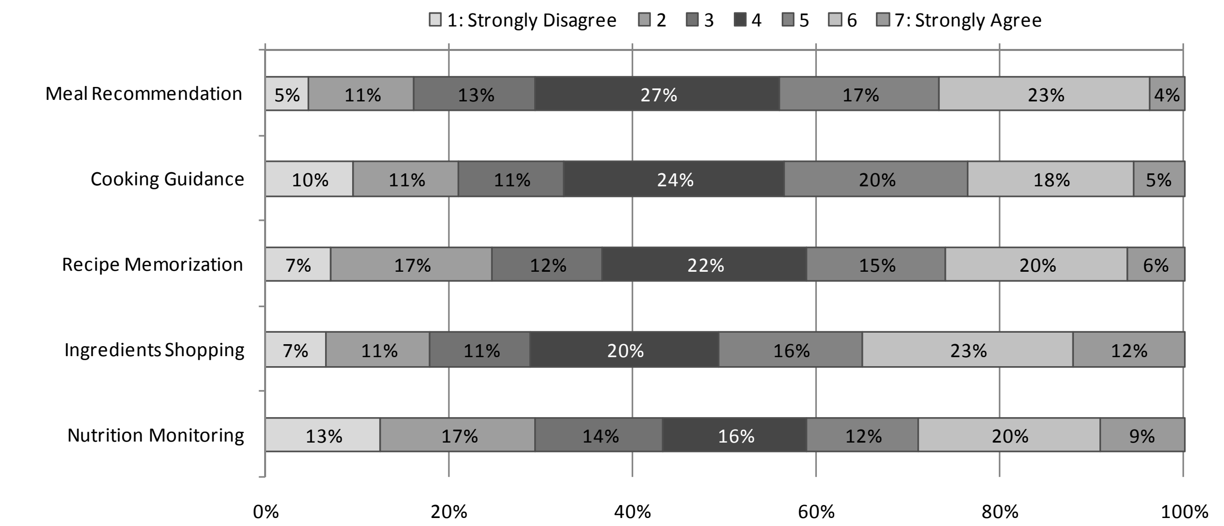
Note: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Next, we asked whether respondents think they would use the different scenarios frequently (Table 29). Only *Ingredients Shopping* was rated significantly positive (51% positive vs. 28% negative answers). *Meal Recommendation, Cooking Guidance,* and *Recipe Memorization* were rated quite neutral with a slightly positive tendency. *Nutri-*

tion Monitoring received more negative (44%) than positive (41%) answers, which supports our finding from the last question that this scenario achieved a comparably weak perception. Comparing the answers with the score levels of other items, respondents seem to be doubtful whether they would frequently use the features offered by the persuasive kitchen environment.

Table 29: Perception of Scenarios: BI2

BI2: I would use the product concept frequently.



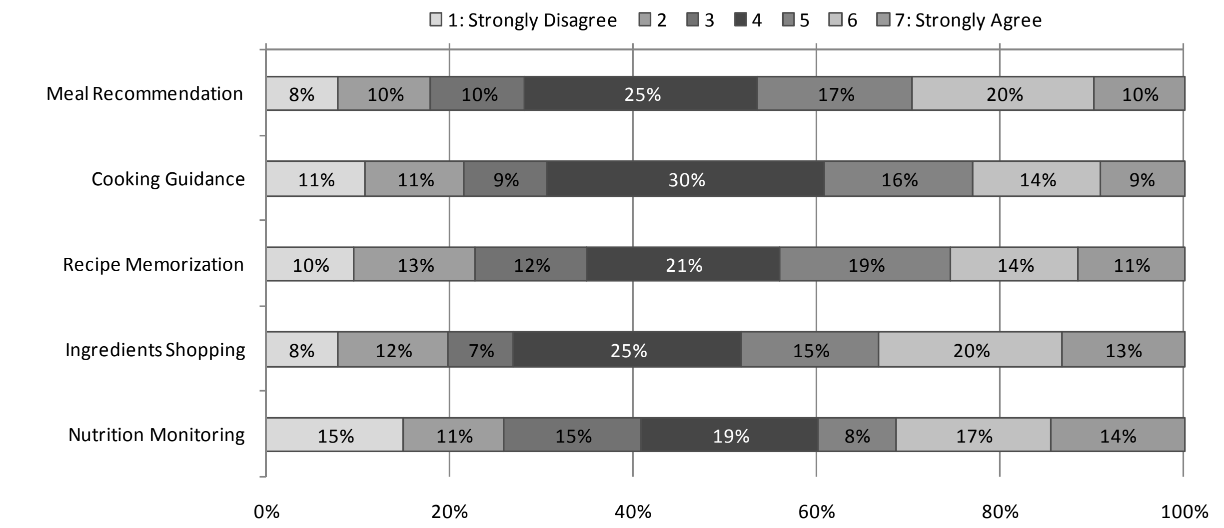
Scenario	Mode	Mean	Std. Dev.	Deviation from scale mean	Sig. (2-tailed)
Meal Recommendation	4	4.23	1.55	0.23	.052
Cooking Guidance	4	4.09	1.69	0.09	.492
Recipe Memorization	4	4.04	1.72	0.04	.753
Ingredients Shopping	6	4.44	1.77	0.44	.002**
Nutrition Monitoring	6	3.93	1.92	-0.07	.658

Note: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Finally, the respondents were asked whether they wished that the proposed concepts were already available. Again all scenarios were rated close to the neutral scale mean, which is reflected by the fact that the most frequent answer for all scenarios was the neutral value of 4. Consistent with the latter two items, *Ingredients Shopping* received the best scoring with 47% vs. 28% negative answers. *Meal Recommendation* was also perceived significantly positive. *Cooking Guidance, Recipe Memorization, and Nutrition Monitoring* were rated close to the neutral level.

Table 30: Perception of Scenarios: BI3

BI3: I wish that the product concept was already available.



Scenario	Mode	Mean	Std. Dev.	Deviation from scale mean	Sig. (2-tailed)
Meal Recommendation	4	4.31	1.72	0.31	.020*
Cooking Guidance	4	4.08	1.74	0.08	.564
Recipe Memorization	4	4.13	1.81	0.13	.348
Ingredients Shopping	4	4.40	1.81	0.40	.005**
Nutrition Monitoring	4	4.04	2.01	0.04	.817

Note: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

To summarize the results for *Behavioral Intention*, there is only a slightly positive attitude towards the scenarios. For all three items, *Meal Recommendation* and *Ingredients Shopping* were ranked best, and *Nutrition Monitoring* consistently received the lowest score.

It is not possible to infer a mapping from these results to a concrete buying decision, since we neither know how other factors (e.g. price) influence this decision nor how the score values correspond to an actual buying decision. However, we can conclude that all scenarios are perceived positive to some degree, and that there is a slight preference for *Meal Recommendation* and *Ingredients Shopping* in comparison to the other scenarios.

Comparing the results for *Behavioral Intention* with the general wish to get assistance (Table 18), respondents expressed a stronger wish for getting assistance than for the intention to use the proposed scenarios. This might be interpreted to mean that some doubts are remaining on whether the proposed kitchen environment can fulfill the respondent's expectations.

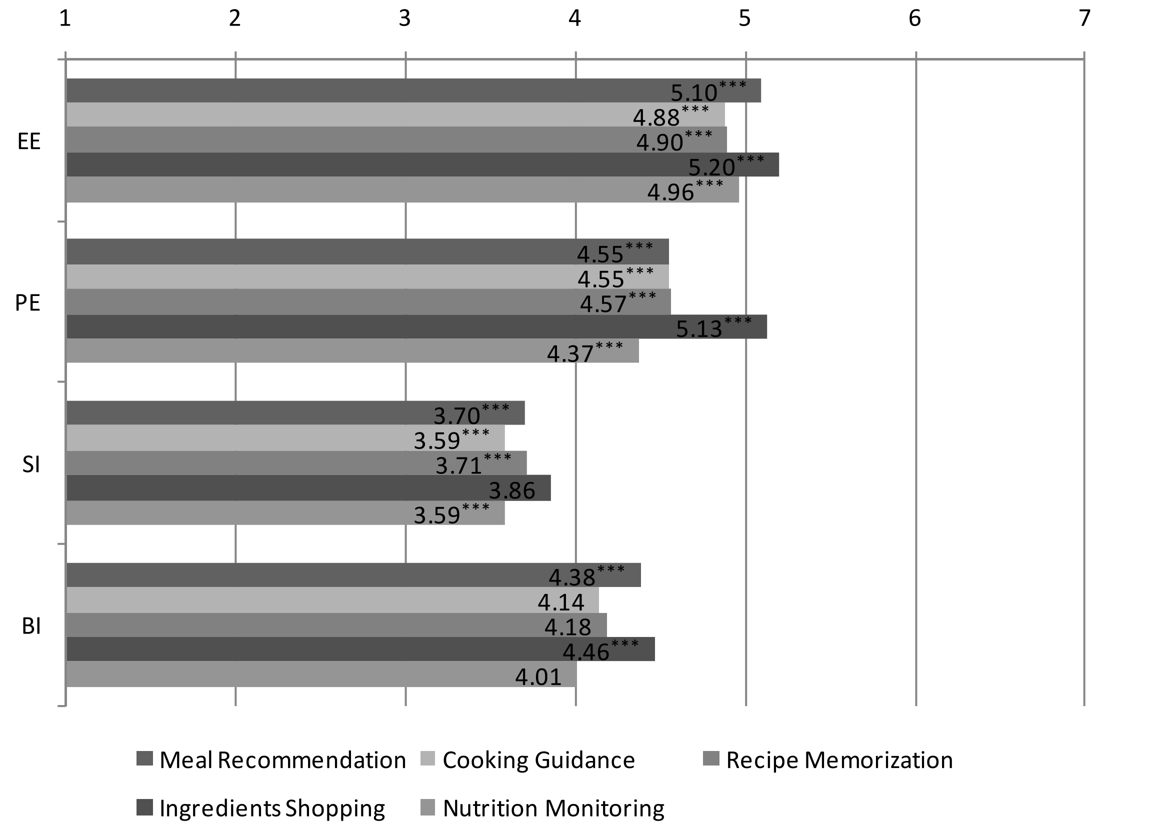
II.5.6.3.5 ***Comparison of scenarios***

To condense and summarize our item level findings, Figure 7 aggregates the survey responses on construct level across the five scenarios. In accordance with the findings presented above, the scenarios have received positive scores for *Effort Expectancy* and *Performance Expectancy*, and negative scores for *Social Influence*. *Behavioral Intention* has been rated neutral with a slightly positive tendency.

Ingredients Shopping has consistently received the best rating across all constructs. *Nutrition Monitoring* has received the weakest rating with respect to *Performance Expectancy, Social Influence,* and *Behavioral Intention*. The other scenarios lie between these extremes. This result supports the conclusion from our respondents' ex-ante statements that hedonic and convenience aspects are more appealing than persuasive features. *Nutrition Monitoring* is most directly linked to supporting users in changing their nutrition habits, but is rated worst among the five scenarios. Evidently, persuasive aspects of the proposed kitchen environment are less appealing than the expectation that food preparation becomes more convenient and joyful.

Comparing the constructs with each other, *Effort Expectancy* has achieved the highest scores, followed by *Performance Expectancy*, *Behavioral Intention*, and *Social Influence.* Evidently, respondents had only few concerns about the usability of the proposed scenarios, and they felt confident that the scenarios will be useful. However with regard to their social environment, they seem to have major concerns about whether their friends and families would appreciate such scenarios. *Social Influence* is the only construct which received in average a negative rating.

Behavioral Intention has been rated comparably weak. The means of this construct are only slightly above the neutral mean, and for three scenarios, it does not significantly differ from the neutral value of 4. As *Effort Expectancy* and *Social Influence* demonstrate that the sample population was willing to give significantly positive and negative ratings, the rating given for *Behavioral Intention* indicates a relatively weak willingness to adopt the proposed scenarios.



(Significance level of deviation from scale mean: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$)

Figure 7: Scenario Comparison on Construct Level

II.5.6.4 ***Correlation Analysis***

We conclude our data analysis with an exploration of correlations between certain demographic factors and the intention to use the proposed persuasive kitchen scenarios. Correlations between *Gender, Age, Personal Innovativeness in IT, Technology-oriented Profession,* and *Number of Children* on the one hand and *Behavioral Intention* (BI) on the other hand are investigated.

Due to the fact that the variables have been measured on different scale levels, different correlation analysis methods are applied (Bortz and Schuster 2010). To analyze whether a significant correlation between the dichotomous variable *Gender* and *BI* exists, a t-test for independent samples is applied. For the remaining variables, it cannot be clearly decided whether they are measured on ordinal or interval scale level because they are either classified (*Age, Number of Children)*, or measured on the basis of a 7-point Likert scale (*PIIT*, *Technology-oriented Profession)*. Therefore we apply Pearson's product-moment correlation coefficient r, and the two rank correlation coefficients Spearman's ρ and Kendall's τ for these variables.

For *Gender*, a t-test reveals that the moderate difference for *Behavioral Intention* is statistically significant (Table 31). So we conclude that men show a higher willingness to adopt the proposed kitchen environment.

Table 31: Correlation Analysis: Gender

Gender	Mean (BI)	Mean Diff.	Sig. (2-tailed)
Male	4.36	0.26	.030*
Female	4.10		
Note: * p < 0.05; ** p < 0.01; *** p < 0.001			

Age does not show a significant correlation with *Behavioral Intention*. *Personal Innovativeness in IT* significantly correlates with *Behavioral Intention* (Table 32). The calculated correlation coefficients lie between 0.135 and 0.183. The degree to which a person has a technology-oriented profession does not significantly correlate with *Behavioral Intention*. The same result was obtained for the number of children a person has.

Table 32: Correlation Analysis

Correlation	Pearson's r	Sig. (2-tailed)	Spearman's ρ	Sig. (2-tailed)	Kendall's τ	Sig. (2-tailed)
Age - BI	-0.009	0.795	-0.021	0.555	-0.016	.542
PIIT - BI	0.183	0.000***	0.182	0.000***	0.135	.000***
Tech. Prof. - BI	0.059	0.090	0.041	0.234	0.033	.218
No. Children - BI	-0.032	0.364	-0.020	0.563	-0.017	.541
Note: * p < 0.05; ** p < 0.01; *** p < 0.001						

II.6 Discussion

II.6.1 Theoretical Implications

Data analysis applying PLS confirmed the applicability of our UTAUT-based research model for the proposed persuasive kitchen environment. The model was able to account for almost 70% of the variance in *Behavioral Intention*, which is equivalent to the results obtained by Venkatesh et al. (2003) for the original model.

Performance Expectancy has shown to have the strongest direct effect on *Behavioral Intention*. *Effort Expectancy* and *Social Influence* act as significant predictors too, but at a weaker level. Comparing our results with Venkatesh et al. (2003), we obtained a similar effect size for *Performance Expectancy* with a path weight of 0.54 (0.53 for the pooled data in the original UTAUT model). The effect of *Effort Expectancy* on *Behavioral Intention* was slightly stronger in our analysis, with a path weight 0.15 versus

0.10 in the UTAUT model. For *Social Influence*, we obtained a much higher path weight of 0.27 compared to the value of 0.02 in UTAUT. The significance of *Effort Expectancy* and *Social Influence* in our setting is underpinned by the strong effect both constructs exert on *Performance Expectancy.* Path weights of 0.36 and 0.48 show that these constructs have a strong indirect effect on *Behavioral Intention,* mediated by *Performance Expectancy*.

Considering the fact that the original work on UTAUT and many subsequent applications of this research model were situated in a professional setting (Schepers and Wetzels 2007; Venkatesh et al. 2003), our findings indicate that the effects of *Effort* Expectancy and *Social Influence* on *Behavioral Intention* are considerably stronger in a private environment than in a professional setting. Evidently, learning efforts and usability aspects are considered to be slightly more important in a home than in an industrial environment.

For *Social Influence* we obtained a similar result. In a majority of technology acceptance studies, the effect of *Social Influence* on *Behavioral Intention* has turned out to be insignificant (Sun and Zhang 2006). When it was found to be significant, the effect was mostly very weak (Venkatesh et al. 2003). In contrast to these previous findings, *Social Influence* exerts a relatively strong influence (path weight 0.27) on *Behavioral Intention* in the investigated kitchen scenario. One reason might be that aspects like image and self-expression play a larger role in the home than in the professional domain, or that consumers perceive the effect of a technology on their personal image to be stronger in the home domain. Another reason may be that a decision for a certain technology at home is not made individually but in cooperation with family members or friends, whereas in an industrial setting, such a decision is rather based on one's own judgment.

Our results further corroborate the findings of earlier analyses that *Effort Expectancy* and *Social Influence* exert a strong effect on *Performance Expectancy* (King and He 2006; Lee et al. 2003; Schepers and Wetzels 2007; Venkatesh and Davis 2000). This result may indicate that consumers do not strictly separate the three aspects but take usability and social aspects into account when assessing the performance of a product.

Our moderator analysis has shown that *Gender* poses a significant moderation on the relationship of *Social Influence* and *Behavioral Intention*. For men, *Social Influence* is a stronger predictor than for women, which can be interpreted that it is relatively more important to men that their friends and families appreciate adopting the proposed technology. The direction of this effect is opposed to what has been found in previous work. Venkatesh and Morris (2000) and Venkatesh et al. (2003) found that *Social Influence* has a stronger effect on *Behavioral Intention* for women than for men. In our home environment setting, this effect is reversed and has a strong effect size. This may indicate that in a domestic setting, decision making follows a different pattern than in a professional environment. Possibly, men feel less certain about their intentions in a home environment than in a professional setting. Previous research has shown that men are more task-oriented (Minton and Schneider 1980) and driven by achievement needs (Hoffman 1972). As tasks and goals are less clearly defined in the proposed kitchen scenario, men might tend to seek advice by friends and family members. Another reason might be that there is a general tendency in families that men leave decisions related to the proposed scenario to their partners, which is then reflected in a stronger influence of *Social Influence*.

In contrast to the original UTAUT model, our data do not support the assumption that *Gender* exerts a significant influence on the other relationships in our study. There was a weak effect that women are more driven by *Effort Expectancy* and *Performance Expectancy* than men, but both effects were insignificant.

The same holds true for *Age*. Weak but insignificant effects could be found for *Effort Expectancy* and *Social Influence*, which indicates that older people pay more attention to *Effort Expectancy*, whereas younger people are more influenced by their friends and families.

Regarding the moderators that were added to the original model, only *Importance* showed a significant effect on the relationships of *Performance Expectancy* and *Social Influence* on *Behavioral Intention*. *Importance* indicates to which degree respondents feel a certain need to get the proposed assistance. We found that with increasing *Importance*, *Performance Expectancy* has a relatively stronger effect on *Behavioral Intention*, whereas the effect size of *Social Influence* decreases. This may be interpreted that people who feel a strong need to get the proposed assistance in their kitchen tend to base their judgment primarily on performance aspects and neglect social aspects. They seem to be driven by their personal need and tend to use the product if they believe that their need will be fulfilled. In contrast, people who do not feel a strong need for the proposed assistance tend to take into account the opinion of others and consequences for their social image.

With *Personal Relevance*, we measured a person's dedication and interest in the analyzed application domain, which means concretely whether a person is interested in cooking. We found that *Personal Relevance* did not exert any moderator effect on one of the relationships in the research model.

Importance and *Personal Relevance* were derived from the often applied construct *Involvement*, which was defined by Barki and Hartwick (1994) as "a subjective psychological state, reflecting the importance and personal relevance of an object or event". We argued that *Importance* and *Personal Relevance* are two very different aspects of a psychological state that should be measured separately.

The result that *Importance* turned out to be a significant moderator, while *Personal Relevance* did not have a significant influence on one of the relationships, supports our theoretical consideration that the often applied construct *Involvement* should be separated into the two aspects of *Importance* and *Personal Relevance,* as they evidently exert a different kind of moderating influence on the relationships in the model.

For *Personal Innovativeness in IT (PIIT),* we expected that higher degrees of *PIIT* would diminish the influence of *Effort Expectancy* and *Social Influence* and increase the influence of *Performance Expectancy*. We assumed that respondents with a high degree of *PIIT* were self-confident about operating novel devices and therefore were not afraid of learning efforts. Furthermore, we expected that IT-affine respondents are to some degree enthusiastic about new technologies and therefore are less influenced by the opinion of others when making a decision about whether to adopt it. At the same time we expected that the relative importance of *Performance Expectancy* would increase because a high level of technical innovativeness might indicate a good understanding for technical capabilities and therefore allows for a more accurate assessment of the technical features of the proposed scenarios. None of our hypotheses could be confirmed. *PIIT* does not exert a moderating influence on the relationships between *Effort Expectancy, Performance Expectancy*, and *Behavioral Intention*. This result is consistent with previous work. We borrowed the *PIIT* construct from Agarwal and Prasad (1998), who also investigated its moderating role in a technology acceptance model. They also found that *PIIT* does not exert a moderating influence on the paths from *Effort Expectancy and Performance Expectancy* to *Behavioral Intention*.
For the link between *Social Influence* and *Behavioral Intention,* we found a significant moderator effect, which is opposed to the assumed direction. Thus we could not confirm the hypothesis that a higher level of *PIIT* reduces the importance of *Social Influence* as a predictor for *Behavioral Intention*, and including this effect does not significantly improve the explained variance of *Behavioral Intention*. Nevertheless, our analysis has shown that there might be a strong influence of *PIIT* on the SI-BI relationship such that tech-savvy people tend to be influenced more than others by their social en-

vironment. This effect may be explained in two different ways. First, it is possible that tech-savvy people indeed rely more on the opinion of their social environment than other people do. This may either be due to certain personality traits specific to tech-savvy people or because they have experienced that their friends and families do not share their enthusiasm for novel technologies. Second, it is also possible that tech-savvy people are less doubtful about the opinion of their social environment. They are enthusiastic about a novel technology and therefore assume that others will share their opinion. At the same time they rate the proposed scenarios better than others. As a consequence, the SI-BI path weight is higher for tech-savvy people than for others, but the underlying reason is that they have a different perception of what others may think.

All considered, the basic structural model could be confirmed, but only few moderator effects could be found. Whereas *Gender* could be approved as a moderator in many studies, *Importance* seems to have a significant, albeit so far underestimated moderating role. In the original UTAUT model and subsequent analyses, *Voluntariness of Use* could be shown to exert a moderating influence. This, however, is usually only applicable in an organizational environment where people can be mandated to adopt a certain technology. In the private domain, adoption decisions are rather made on the basis of individual and voluntary considerations. Therefore, it seems appropriate to complement *Voluntariness of Use,* which measures extrinsic pressure, by a construct that corresponds to factors such as intrinsic pressure to act or personal need. Against this background, we introduced *Importance* as a measure of this psychological state, in order to reflect the intrinsic motivation to accept a novel technology in the private domain. Our results have shown that *Importance* should be further analyzed with regard to its moderating role. In addition to that, *Personal Innovativeness in IT* turned out to have a counter-intuitive effect on the role of *Social Influence*, which should also be further investigated in future research effort.

Comparing our results across the 5 different scenarios, we have seen that the direct effects in the structural model are quite robust. In all five scenarios (and consequently in the aggregated scenario), *Performance Expectancy* exerts the highest influence on *Behavioral Intention*, followed by *Social Influence* and *Effort Expectancy*. Although the path weights themselves were significantly different between some scenarios, the order of effect sizes remained stable.

The results obtained in our moderator analysis were less stable across the scenarios. *Gender, Importance*, and *PIIT* exerted significant moderator effects in the aggregated scenario. Analyzing each scenario individually, *Gender* and *PIIT* were insignificant in all sub-scenarios with regard to the SI-BI relationship, and the moderating influence of

Importance was significant in only two scenarios. One explanation might be that multicollinearity leads to statistical distortions. This assumption is supported by high Variance Inflation Factors and the fact that regression coefficients of the interaction effects are extremely high, in one case even above 1. Another explanation might be that the moderator effects are indeed not stable across the different scenarios and should be interpreted with caution. Whereas our results clearly indicate that some moderator effects exist in the proposed scenario, generalizability to other scenarios and domains remains questionable due to a lack of robustness of these effects.

II.6.2 Practical Implications

Besides the aforementioned theoretical implications, our study also allows for drawing conclusions relevant to management practice, particularly in the home appliance industry. Our study has shown that cooking plays an important role for our respondents and that they wish to have technical assistance for preparing so far unknown recipes.

Descriptive results indicate that our respondents rated the proposed scenarios quite neutral with a slightly positive tendency. Values for the *Behavioral Intention* construct were found to lie between 4.01 and 4.46 (Figure 7). On the one hand, this result shows that the respondents were in average undetermined whether they would adopt the proposed scenarios. On the other hand, about 40% of the respondents gave a positive answer to questions related to *Behavioral Intention*. To interpret this result from a practical perspective, it may indicate that the proposed scenarios will be perceived positive by a large part of the population, although the majority seems to remain doubtful.

The result is robust across different population groups. Gender, age, number of children, and technology-related professional background were not significantly correlated with *Behavioral* Intention. Only the degree to which a person is innovative with respect to technologies had a significant positive correlation with *Behavioral Intention*, which may indicate that such people might be a target group for the early adoption phase.

The ex-ante desire for assistance has been rated much higher (about 60% positive answers) than the intention to use the product (about 40% positive answers). Although other factors not considered in this study might be attributable for this effect, this discrepancy may indicate that the respondents were doubtful about whether the scenarios can fulfill their expectations.

We can further learn from this empirical investigation that consumers regard functional capabilities of a persuasive kitchen environment as key to their adoption decision, whereas usability aspects play only a minor role.

Rather critical were the results with respect to *Social Influence*. A majority of our respondents had concerns about how their social environment would perceive a potential adoption of the proposed scenarios. Our structural model analysis has furthermore shown that the opinion of friends and family might be an important determinant of *Behavioral Intention.* This outcome is even more remarkable as past technology acceptance studies have mostly found only a weak or insignificant relationship between *Social Influence* and *Behavioral Intention*. Surprisingly, the degree to which *Social Influence* affects *Behavioral Intention* is different for men and women such that men are influenced more strongly by their social environment than women.

From a practical perspective, this result might indicate that *Social Influence* has so far been underestimated and should be taken into account more thoroughly in the planning phase of a product innovation in the home equipment domain. Moreover, our results may contradict past convictions that women are more influenced by their social environment than men. In this specific domain, marketing activities should consider the comparably stronger adherence of men to the opinion of their social environment.

Our analysis further indicates that our sample population falls apart into two segments. One segment stated ex-ante (i.e. before knowing the scenarios) that they wish to have the proposed assistance. This segment seems to base its behavioral intention rather on performance considerations than on the opinion of others. The other segment has shown a lower ex-ante need for the proposed assistance. This segment is influenced more strongly by their social influence, and is driven to a lesser degree by performance expectations.

Another potential customer segmentation is determined by *Personal Innovativeness in IT*. Not only does this character trait directly correlate with *Behavioral Intention*, but also influences to which degree *Behavioral Intention* is determined by *Social Influence*. Our results show that, against our expectations, more innovative persons are more influenced by their social environment. We would have expected that these people are more motivated by technical features and performance expectations because they are familiar with novel technologies and therefore are capable of making their own judgment. We can only speculate what the underlying psychological mechanisms are. Maybe such people want to avoid a "nerd"-image or feel less self-confident about their decisions. For managerial practice however, we can conclude that particularly those people, for whom a novel technology is most appealing, and who are a formidable target group for early adoption, may have extraordinary concerns with respect to their image and the opinion of their social environment.

Comparing the five proposed scenarios, *Ingredients Shopping* - i.e. the generation of a shopping list for a certain recipe in consideration of already available ingredients - has been rated best, followed by *Meal Recommendation. Cooking Guidance, Recipe Memorization*, and *Nutrition Monitoring* have been rated slightly worse. Although differences in the acceptance levels of the five scenarios are small, *Nutrition Monitoring* has gained the least attraction. Furthermore, a majority of our respondents have stated that healthiness of a meal is less important for them than tastiness. This indicates that convenience features and hedonic value may be a more appealing buying argument for potential customers than persuasive features. Although support for achieving nutritional goals is a welcomed side effect, it is evidently not the feature that is most appreciated by our respondents.

Attention should be paid to the social aspects of such technologies. As could be shown, the opinion of friends and families may be important for an adoption decision, and our sample population showed some doubts about the social acceptance of adopting the proposed scenarios. Consequently, marketing measures should not only focus on technological capabilities. In addition, an image campaign seems to be necessary such that potential adopters get the feeling that they improve their social image, or at least do not harm it, when they use a persuasive kitchen environment at home. This aspect is of particular importance for appealing male customers, technology enthusiasts, and customers who perceive a relatively weak ex-ante desire for kitchen assistance. In contrast to that, females, non-innovators, and people with a strong ex-ante desire for kitchen assistance are more appealed by performance aspects such as technical features and functionalities.

II.6.3 Limitations

Even though every effort has been made to ensure the validity of our findings, the present study comes with limitations that point to opportunities for further research. First of all, while the size of our sample is sufficient for testing the proposed structural model, larger samples would be helpful to investigate simultaneously the differences in adoption behavior between geographic regions and additional demographic factors such as income, family status, etc.

Second, although having achieved sufficient explanatory power, our results nevertheless leave room for additional factors not included in our research model that might influence adoption behavior. We therefore propose to discuss and empirically test the relevance of other constructs beyond the scope of the present study.

Third, our investigation has been based on scenario descriptions, which limits the transferability to a commercial offering. As a consequence, the scenarios should next be implemented and tested in an experimental setting to further assess the validity of our results.

Forth, the counter-intuitive interaction effects caused by *Gender* and *Personal Innovativeness in IT* should further be investigated with regard to generalizability and robustness. Furthermore, the moderating role of *Importance,* which was newly introduced and confirmed in this study, requires additional empirical studies in order to determine whether it is a valuable contribution to technology acceptance research.

III Persuasion Profiling

III.1 Case Description

The goal of the study presented in this chapter is to investigate the second research question raised in this dissertation:

Q2: *Can persuasion profiling improve the effectiveness of persuasive messages?*

Many prototypical implementations of persuasive technologies have been proposed to achieve behavioral changes - for example, aiming at physical activity (Consolvo et al. 2009; Lacroix et al. 2009; Lin et al. 2006), a healthy diet (Lo et al. 2007; Peng 2009), losing weight (Arteaga et al. 2009), stop smoking (Grolleman et al. 2006), saving energy (Loock et al. 2011; Shiraishi et al. 2009), or adopting fuel-efficient transportation habits (Froehlich et al. 2009; Lee et al. 2010; Meschtscherjakov et al. 2009). These examples have in common that they apply a static set of persuasive strategies but do not adapt the way a person is influenced to his personality traits.

Recently, persuasive systems have come into the focus of IS research that adapt their persuasive strategies to the personality traits of a user. Psychological research has shown that persons respond differently to certain persuasive principles. For example one person may rather be influenced by personal goal setting whereas another person may rather be susceptible to social norms. So-called *adaptive persuasive technologies* can implement different persuasive principles and select the one that is most promising for a certain user. Following a proposal of (Kaptein and Eckles 2010a), we define *adaptive persuasive technologies* as "technologies that aim to increase the effectiveness of some attitude or behavior change by responding to the behavior of (and other information about) their individual users".

Adaptive persuasive technologies must be capable of *persuasion profiling*, i.e. to retrieve *persuasion profiles,* which describe the expected susceptibility of an individual to different persuasive strategies. *Persuasion profiles* are defined by (Kaptein and Eckles 2010a) as "collections of expected effects of different influence strategies for a specific individual". Two alternative approaches can be followed to obtain persuasion profiles. Either the persuasive profile is heuristically determined by demographic factors (e.g. gender, age, education level), or it is measured for each individual specifically. Persuasion profiles can be measured ex-ante for example by exposing users to a questionnaire before the system is used. Alternatively a learning approach can be applied by evaluating the effectiveness of different persuasive principles over time.

In the present study, we investigate to which degree *persuasion profiling* can increase the effectiveness of adaptive persuasive systems. For this purpose, an experiment was conducted in which subjects were exposed to persuasive SMS messages under three experimental conditions. One group received messages that fit to their personality traits, a second group obtained messages that do not fit, and a third group was exposed to a random selection of messages. Comparing the degree to which the three experimental groups responded to the messages, we could show that well-fitting messages and randomly selected messages perform significantly better than non-fitting messages, whereas the difference between well-fitting and randomly selected messages was not significant.

The following sections present the details of this experiment. In the next section, we summarize work that is related to the profiling of persuasive technologies. Then the research design is described, followed by an explanation of the data collection process. Next, the data analysis method (based on linear mixed models) is described, followed by a discussion of theoretical and managerial implications.

III.2 Related Work

The central question investigated in this survey is whether persuasion profiling can leverage the effectiveness of adaptive persuasive technologies. Various authors in the fields of social psychology and information systems research have investigated the effectiveness of different persuasive strategies. We structure the different streams of related work along three dimensions. The first work stream investigates which persuasive principles are prevailing, the second work stream contrasts the effectiveness of different strategies against each other, and the third work stream, which is most closely related to the research question underlying this study, evaluates how personality traits relate to the susceptibility to certain persuasive strategies.

III.2.1 Persuasive Strategies

Several taxonomies have been developed over the last decades to structure the potential approaches to exert persuasion on a person. Among the first of them, Marwell and Schmitt (1967) identified 16 basic persuasive strategies, which they clustered into the five groups of *rewarding activities*, *punishing activities, expertise, activation of impersonal commitment,* and *activation of personal commitments.* Levine and Wheeless (1990) compiled a list of 53 basic persuasive strategies, which were derived from nine earlier taxonomies. Kellermann and Cole (1994) analyzed 74 classification systems and developed a taxonomy of 64 persuasive principles. With a focus on persuasive technologies, Fogg (2002) developed a taxonomy of 42 persuasive strategies, which are clustered along six functional aspects of IT-based persuasive systems.

Among the many available taxonomies, we decided to apply the rather parsimonious taxonomy proposed by Cialdini (2008). It reduces the vast number of persuasive tactics to six clearly distinguished strategies:

(1) Authority
People have a tendency to obey directives from authorities. This tendency may persist even if the directive is contradictive to the actual intention or attitude of a person. In some cases, authority is justified by true expertise, experience or legitimate power. However, the automatic obedience to authorities, which is deeply rooted in society, often makes people react to symbols of authority (e.g. titles, clothes or vehicles) instead of questioning the justification of authoritative behavior.

(2) Commitment and Consistency
People desire to stay consistent with their attitudes, convictions, and particularly with their commitments. If a person has committed to a certain attitude, he will tend to follow a request if it is consistent with earlier commitment. Even commitments that turned out to be disadvantageous will often persist due to the strong desire to stay consistent.

(3) Social Proof
Confronted with an uncertain situation in which people do not know how to behave, people tend to act as they think similar others would do. As a persuasive strategy, this tendency can be exploited by making a person believe that the requested behavior is consistent with the attitude of similar others. The strategy is most effective if the situation is uncertain or ambiguous, and if the reference group is similar.

(4) Liking
People have a higher willingness to comply with a persuasive attempt from a person that they know or like. The degree to which a person is liked is influenced by several factors: Physical attractiveness, similarity, praise and compliments, familiarity, and positive associations.

(5) Reciprocity
People desire to return a favor they received from another person. Doing a favor in advance to a persuasive attempt makes a person feel obliged to return this favor. This strategy remains effective even if the favor was not requested. Fur-

thermore, it may provoke a person to return a greater favor than was originally received.

(6) Scarcity

People value scarce items or opportunities more than broadly available ones. This is often exploited as a persuasive strategy by restricting the time for making a decision, or by suggesting scarce availability. The principle is not limited to physical goods, but also extends to information. Information that is difficult to obtain exerts more persuasive power than broadly available, public information.

The six persuasive strategies are often applied, for example, in sales and marketing to influence buying decisions. They also build the basis for bargaining strategies. Table 33 summarizes the six persuasive principles and indicates potential ways to exploit them for persuasive intentions (Cialdini 2001).

Table 33: Cialdini's Persuasive Strategies

Principle	Description	Exploitation
Authority	People defer to experts.	Exposition of expertise; use of symbols that indicate authority.
Commitment and Consistency	People align with their clear commitments.	Demand for active, public, and voluntary commitments; reference to prior commitments; start with small request, then advance with a larger request (foot-in-the-door tactic).
Social Proof	People follow the lead of similar others.	Reference to the behavior of similar peers.
Liking	People like those who like them.	Emphasis on similarities; genuine praise.
Reciprocity	People repay in kind.	Doing favors in expectation of a return.
Scarcity	People want more of what they can have less of.	Highlighting of unique benefits and exclusive information; setting scarce timelines; indicating limited availability.

III.2.2 Influencing Factors for the Effectiveness of Persuasive Strategies

Normative social influence describes the tendency to comply with expectations, actions or attitudes of others in order to be liked and accepted. Consequently, it describes the psychological mechanisms underlying the persuasive strategy of *social proof*. Numerous studies have shown the powerful effect of normative social influence on human behavior when individuals observe the actions of others (for a review see Cialdini and Goldstein (2004)). More recently, researchers investigated how different implementations of normative social influence affect its effectiveness. Schultz (1999), Goldstein et al. (2008) and Haines and Spear (1996) have shown that social norms do not need to be observed directly to become effective but can also be conveyed in written form. Source credibility plays a major role for the effectiveness of persuasive messages. It describes the perceived expertise and trustworthiness of a message source (Kelman and Hovland 1953). Tormala et al. (2006) have shown that high source credibility can leverage the effect of persuasive messages that generate primarily positive thoughts, but that the effect is reverse if the primary perception of a message is negative. In their *Theory of Normative Conduct*, Cialdini et al. (1991) distinguish between *descriptive* and *injunctive* norms. Descriptive norms convey information about what others really do, whereas injunctive norms inform about what others approve or disapprove. To become effective, both kinds of social norms must be focal in attention and salient in consciousness. Cialdini et al. (2006) have further demonstrated that descriptive persuasive messages must be formulated carefully. Complaining about a regrettable status-quo may have a counterproductive effect because it may establish a social norm of the undesired behavior. In their study, descriptive persuasive messages complaining about theft of petrified wood in a national park lead to an increase of thefts, whereas injunctive feedback reduced theft rates.

This effect was confirmed by Loock et al. (2011), who analyzed how descriptive and injunctive normative feedback affects energy consumption. Whereas for below-average consumers, descriptive feedback increased energy consumption (as it established the social norm of higher consumption), above-average consumers adjusted their behavior such that they reduced energy consumption. A combination of injunctive and descriptive feedback decreased consumption in both groups.

Nolan et al. (2008) investigated whether normative social feedback is more effective in promoting energy-saving behavior than factual feedback (i.e. neutral information about energy consumption or costs). They found that social feedback was superior to factual feedback although study participants reported that they found normative messages less motivating. Similarly, Midden and Ham (2009) confirmed this finding and further demonstrated that negative feedback (both factual and social) has a stronger effect than positive feedback. They conclude that the strongest effect can be achieved by giving negative social feedback.

The studies outlined in this section are similar in that they analyze the effectiveness of different persuasive approaches without taking into account individual differences between subjects. In the following section, we review related work that analyzes the impact of individualized persuasive interventions.

III.2.3 Influence of Individual Differences

Noar et al. (2007) conducted a meta-analytic review of 57 health-related intervention studies to evaluate whether tailored persuasive messages are superior to non-tailored ones. *Tailoring* is defined as "any combination of strategies and information intended to reach one specific person, based on characteristics that are unique to that person, related to the outcome of interest, and derived from individual assessment" (Kreuter et al. 2000). They found that tailored messages perform generally better than non-tailored messages. They furthermore found that tailoring is most effective if it is based on theoretical concepts and personality traits like attitudes, self-efficacy, stage of change, processes of change, and social support. In contrast, tailoring along demographic characteristics like gender, age, education level or racial and ethnical groups cannot significantly improve effectiveness. With regard to cultural differences, Cialdini et al. (1999) found that cultural conditioning may influence individual susceptibility to certain persuasive strategies. They compared the effectiveness of *social proof* and *commitment / consistency* messages in the United States and Poland. Their results have shown that *social proof* performed better in the more collectivistic culture in Poland, whereas *commitment / consistency* was superior in the more individualistic society of the United States.

Analyzing individual differences in the processing of persuasive messages, Cacioppo et al. (1986) found that people with a high *Need for Cognition* (NfC) think more intensively about incoming messages than people with low NfC. In order to avoid cognitive elaboration, low NfCs evaluate persuasive messages rather heuristically than cognitively. High NfCs show stronger compliance with their attitude than low NfCs. At the same time, high NfCs are influenced by convincing arguments, whereas low NfCs tend to be susceptible to persuasive tactics. These findings explain - at least partially - the results found by Kaptein et al. (2009) and Kaptein et al. (2010) that people differ in their general level of susceptibility to persuasion.

Moon (2002) investigated how personality traits influence the effectiveness of different persuasive strategies. They compared the response of dominant and submissive persons to dominant and submissive message styles. They found that neither dominant nor submissive messages are generally superior. In fact, dominant personalities are more susceptible to dominant messages, whereas submissive messages show a larger effect when applied to submissive personalities. Similarly, Halko and Kientz (2010) have shown that the perception of differently shaped persuasive approaches is influenced by the so-called Big-5 personality traits, which are *Neuroticism, Conscientiousness, Agreeableness, Extraversion, and Openness* (Goldberg 1993).

A widely applied implementation of the *commitment / consistency* strategy is the foot-in-the-door tactic. A small request is made to a person to create intrinsic commitment. Once the person has agreed, the actually intended larger request is revealed. As the person still feels committed, he may also accommodate the larger request to avoid cognitive dissonance and stay consistent with prior attitudes. Cialdini et al. (1995) have shown that this tactic is only effective for individuals that have a high *Preference for Consistency* (PFC). They furthermore found that only half of their study participants showed a high PFC level. Besides that, the study demonstrated that it is possible to measure the susceptibility to a certain persuasive strategy such that it becomes possible to predict the effectiveness of its implementation. Guadagno et al. (2001) have demonstrated that an explicit reference to prior commitment increases the compliance of people with a high PFC level, whereas it has a reverse effect on people with low PFC. The study also confirmed the finding from Cialdini et al. (1995) that people with a low PFC show only a low level of compliance with commitment-based persuasive interventions.

To summarize, there is broad empirical evidence that people differ in their general susceptibility to persuasive attempts as well as in their response to certain persuasive principles. Several studies indicate that applying an inappropriate strategy may reverse the intended effect such that a subject might not only deny compliance with a persuasive message, but might even show an adverse change in behavior. Preferences cannot usually be predicted on the basis of demographic characteristics. Susceptibility for certain strategies is to some degree related to personality traits like dominance or submissiveness, but for selecting an optimal persuasive strategy, individual susceptibility should be assessed.

III.3 Research Design

III.3.1 Experimental Design

To investigate whether profiling may increase the effectiveness of persuasion, an experiment was designed, in which subjects were exposed to persuasive SMS messages that aimed at motivating them to adopt a healthier nutrition behavior by reducing the number of snacks taken in during a day. Subjects were asked to keep a nutrition diary over two weeks. In the first week, no intervention took place to obtain baseline information about the individual nutrition behavior of each subject (baseline phase). In the second week, an SMS message with persuasive content was sent daily to each subject (treatment phase).

Before the start of the experiment, the subjects' responsiveness to the different persuasive strategies was assessed by a personality questionnaire. As a result of this questionnaire, we obtained a susceptibility score for each persuasive principle so that the best- and worst-fitting persuasive principle could be identified for each subject.

Subjects were assigned randomly to three experimental groups. One group was exposed to messages that fit to their personality traits ("RIGHT" condition), another group received messages that do not fit ("WRONG" condition), and a third group obtained randomly selected messages ("RANDOM" condition). The effectiveness of the different treatment conditions was evaluated by comparing the degree of behavior change from the baseline to the treatment phase across the different experimental groups. Figure 8 illustrates the research process followed in this study.

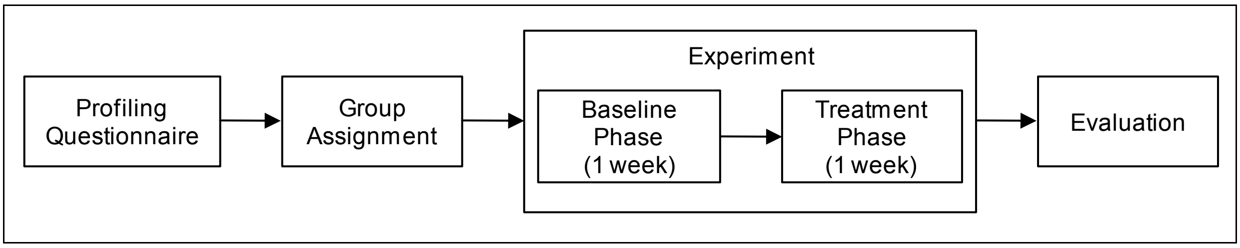


Figure 8: Research Process

Nutrition behavior was monitored via a web-based nutrition diary. Subjects were asked to enter the number of unhealthy snacks taken in over the day into a web form each evening. As daily data entries could not be enforced, the number of measurements varied between subjects. The independent variables are *condition* (WRONG, RIGHT, RANDOM), and phase (1, 2). The dependent variable is the number of unhealthy snacks (SNACKS) taken in during the day. The resulting experimental design can be represented as a four-level hierarchical data structure. Level 1 data represents the repeated measures of the dependent variable SNACKS. The second level describes the experimental phase (1: baseline; 2: treatment). The third level contains the subjects, which are clustered under the three experimental conditions on the fourth level (WRONG, RIGHT, RANDOM). Figure 9 illustrates the hierarchical data structure obtained from the experiment.

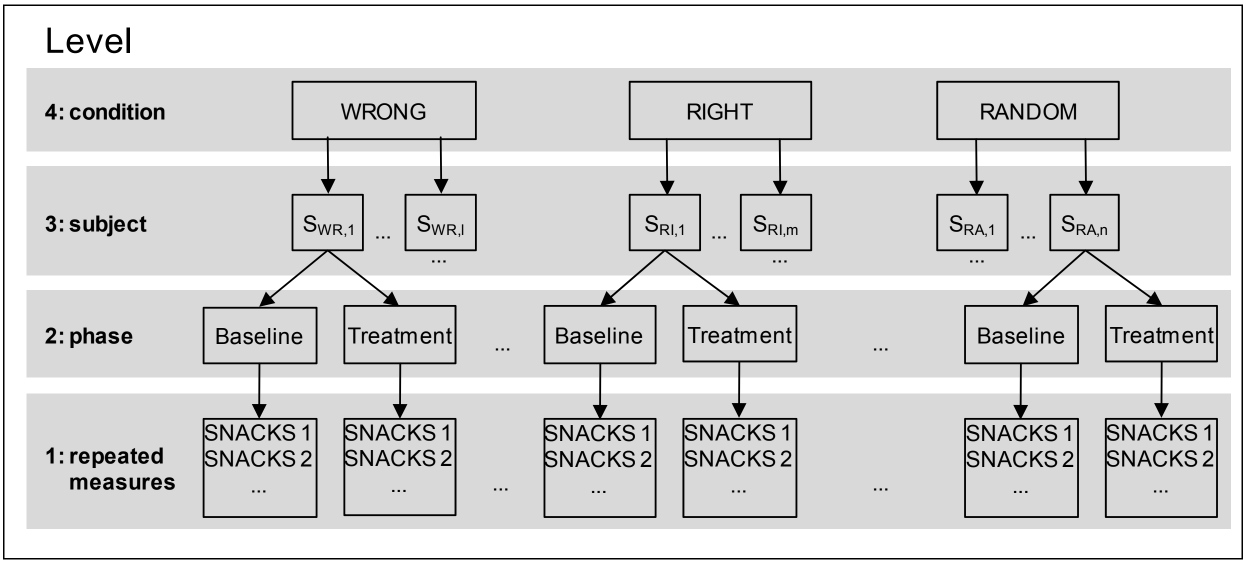


Figure 9: Hierarchical Data Structure

III.3.2 Hypotheses

The purpose of this study is to assess whether adapting persuasive interventions to the individual's responsiveness to different persuasive principles may increase the effectiveness of the interventions. In the following, hypotheses are formulated that will be tested on the basis of the data obtained from the experiment outlined above.

Before comparing the different experimental conditions, we will test the hypothesis that there is a significant influence of the persuasive messages under each condition. We formulate the following hypotheses:

> **H1a:** The number of SNACKS consumed per day is lower in the TREATMENT phase than in the BASELINE phase under the WRONG condition.
>
> **H1b:** The number of SNACKS consumed per day is lower in the TREATMENT phase than in the BASELINE phase under the RIGHT condition.
>
> **H1c:** The number of SNACKS consumed per day is lower in the TREATMENT phase than in the BASELINE phase under the RANDOM condition.

We further theorize that under the WRONG condition, the reduction of the number of snacks is lower than under the RIGHT condition. We further expect that the results for the RANDOM condition lie between the WRONG and RIGHT condition. We formulate the following hypotheses:

H2: Under the WRONG condition, the decrease in the number of SNACKS from the baseline to the treatment phase is lower than under the RANDOM condition.

H3: Under the RIGHT condition, the decrease in the number of SNACKS from the baseline to the treatment phase is higher than under the RANDOM condition.

H4: Under the WRONG condition, the decrease in the number of SNACKS from the baseline to the treatment phase is lower than under the RIGHT condition.

III.4 Data Collection

Before starting with the actual experiment, the instruments applied in this study were evaluated. This section describes the development and validation of the profiling questionnaire and the persuasive messages. Furthermore, it explains the sample selection process and the experimental data collection.

III.4.1 Instrument Development: Profiling Questionnaire

Before the start of the actual experiment, subjects were asked to fill in a profiling questionnaire. The purpose of this questionnaire was to assess which persuasive principle each subject would be most susceptible to. On the basis of this assessment, the persuasive messages sent to each subject were selected. Depending on the randomly assigned condition, subjects received either a fitting, a non-fitting, or a randomly selected message. The sample size for this profiling questionnaire was N=333.

Cialdini formulated six persuasive principles: Authority, Commitment, Social Proof, Liking, Reciprocity, and Scarcity (Cialdini 2001, 2008). The profiling questionnaire contained several items associated with each of these principles, which constitute the latent constructs measured by the questionnaire. The items and constructs were formulated in accordance with Kaptein et al. (2009) and Kaptein and Eckles (2010b) and measured on a 7-point Likert scale (Table 34).

Table 34: Measurement Instruments: Profiling Questionnaire

Contruct	Item	Question
Authority	A1	I always follow advice from my doctor.
	A2	I always obey directions from my superiors.
	A3	I am more inclined to listen to an authority figure than a peer.
	A4	I am very inclined to listen to authority figures.
	A5	I'm more likely to do something if told, than when asked.
	A6	When a professor tells me something I tend to believe it is true.
Commitment	**Comm1**	Whenever I commit to an appointment I always follow through.
	Comm2	I try to do everything I have promised to do.
	Comm3	When I make plans I commit to them by writing them down.
	Comm4	Telling friends about my future plans helps me to carry them out.
	Comm5	Once I have committed to do something I will surely do it.
	Comm6	If I miss an appointment, I always make it up.
Social Proof	SP1	If someone from my social network notifies me about a good book, I tend to read it.
	SP2	When I am in a new situation I look at others to see what I should do.
	SP3	I will do something as long as I know there are others doing it too.
	SP4	I often rely on other people to know what I should do.
	SP5	It is important to me to fit in.
Liking	**L1**	I accept advice from my social network.
	L2	When I like someone, I am more inclined to believe him or her.
	L3	I will do a favor for people that I like.
	L4	The opinions of friends are more important than the opinions of others.
	L5	If I am unsure, I will usually side with someone I like.
Reciprocity	**R1**	I always pay back a favor.
	R2	If someone does something for me, I try to do something of similar value to repay the favor.
	R3	When a family member does me a favor, I am very inclined to return this favor.
	R4	When I receive a gift, I feel obliged to return a gift.
	R5	When someone helps me with my work, I try to pay them back.
Scarcity	**S1**	I believe rare products (scarce) are more valuable than mass products.
	S2	When I want to buy a product, I feel happy if I grab the last item of this product in the shop.
	S3	Products that are hard to get represent a special value.
	S4	When my favorite shampoo is almost out of stock I buy two bottles.
	S5	When my favorite shop is about to close, I would visit it since it is my last chance.
Items marked bold were applied after Exploratory Factor Analysis.		
Source: Kaptein et al. (2009); Kaptein and Eckles (2010b)		

Reliability and validity of the questionnaire were assessed by Exploratory Factor Analysis (EFA). The original questionnaire showed insufficient reliability in terms of low Cronbach's α values and indicator loadings. Subsequently, the following items were eliminated: A1, A3, A5, Comm3, Comm4, SP1, L4, L5, S4, S5. The resulting reduced item set has been applied for the profiling of the subjects. In the following, EFA results for the final item set are summarized.

Principal Component Analysis (PCA) with oblique rotation (direct oblimin, delta = 0) was applied as the EFA procedure. Kaptein et al. (2009) have shown that people differ in their general susceptibility to persuasion. Therefore it can be expected that high scores for one principle are typically associated with relatively high scores for other principles, which means that the measurements of the susceptibility to individual persuasive principles are correlated with each other. Under such circumstances, oblique rotation, which does not assume uncorrelated factors, is usually recommended and preferred over orthogonal rotation, which in contrast assumes that factors are uncorrelated (Conway and Huffcutt 2003; Brown 2006).

A Kaiser-Meyer-Olkin (KMO) measure of 0.819 and individual MSA (Measure of Sampling Adequacy) values above 0.621 verify the sampling adequacy for the analysis (Hutcheson and Sofroniou 1999). Bartlett's test for sphericity ($\chi^2(231)=2408.993$; $p<0.001$) indicates that item correlations are large enough for the application of PCA. Communalities are well above the recommended threshold of 0.5 for all items except S2.

The determinant of the inter-item correlation matrix is 0.001, which is above the necessary level of 0.00001. Haitovsky's test (Haitovsky 1969) confirms that the determinant is not significantly different from zero ($\chi^2_H=0.323995$; df=231; p=1.000), which indicates that problematic multicollinearity is not present in the given sample.

PCA extracted six factors with Eigenvalues larger than 1 (Kaiser's criterion), explaining 62.1% of the variance. The pattern matrix (Table 35) confirms the assumed factor structure. To decide about the significance of the factor loadings, the sample size has to be taken under consideration. Stevens (2002) recommends considering factor loadings above 0.4 as substantial and above 0.298 as significant for a sample of 300. All factor loadings are well above the threshold of 0.4, and the majority of factor loadings are above or close to the often recommended value of 0.7 (Nunnally and Bernstein 1994). In respect thereof, the factor loading of S2 stands out with a comparably low loading of 0.432. However, S2 has been kept in the final item set despite its low factor loading and communality since the loading is above 0.4, its content validity seems to be ensured, and loadings on other factors are weak. Furthermore, the pattern matrix shows that all cross-loadings are below the significance threshold of 0.298, so that problematic cross-loadings cannot be observed.

As a reliability measure, the Cronbach's α values are above the recommended threshold of 0.7 (Nunnally and Bernstein 1994) for all factors except *scarcity*, which achieves only a value of 0.61. Although being critical, it should be noted that cut-off values act rather as a rule-of-thumb than a sharp demarcation, which is reflected in the fact that consensus about valid cut-off values is still lacking. For example, Robinson et al. 1991 regard a threshold of 0.6 for exploratory research stages as acceptable, and Kline (1999) notes that for psychological constructs, values below 0.7 can also be valid.

Table 35: Pattern Matrix

Item	Factor					
	Authority	**Commitment**	**Social Proof**	**Liking**	**Reciprocity**	**Scarcity**
A2	**.821**	.163	-.079	-.035	-.046	.073
A4	**.810**	-.098	.077	.007	.064	-.006
A6	**.601**	.027	.071	.224	.121	-.050
Comm1	.101	**.720**	-.053	.117	.053	-.005
Comm2	.072	**.703**	-.111	.209	-.043	.096
Comm5	-.054	**.739**	.061	-.128	.057	.012
Comm6	.021	**.740**	.110	-.047	.092	-.060
SP2	.001	.160	**.679**	.088	.021	.026
SP3	.097	.030	**.762**	-.112	-.085	.173
SP4	.045	-.129	**.777**	-.004	.002	.004
SP5	-.099	.007	**.750**	.160	.066	-.099
L1	.176	-.038	.084	**.763**	.008	-.078
L2	.017	-.063	.087	**.781**	.009	.146
L3	-.079	.223	.024	**.706**	.037	-.003
R1	.007	.052	.011	-.095	**.836**	.002
R2	-.080	.121	.059	.091	**.643**	.084
R3	.022	.214	-.048	-.013	**.660**	-.086
R4	.219	-.095	.114	-.145	**.725**	-.015
R5	-.046	-.076	-.096	.206	**.740**	.129
S1	-.090	.039	-.051	-.027	-.004	**.896**
S2	.188	-.011	.126	-.060	.008	**.432**
S3	-.013	-.018	.021	.124	.096	**.796**
Cronbach's α	0.71	0.76	0.76	0.74	0.81	0.61

The correlation matrix (Table 36) shows the correlation coefficients between the factors extracted by PCA. As assumed, the matrix confirms that our factors are notably correlated among each other. Theoretical considerations justify this effect because people differ in their general susceptibility for persuasive interventions. Furthermore, it confirms the ex-ante assumption that oblique rotation will be superior to orthogonal rotation. Inspecting the matrix in detail, several correlations stand out. A*uthority* and *social proof* are highly correlated. From a semantic perspective, both principles have in common that they indicate a tendency to comply with expectations, either from an authority or from social norms. The strong correlation between *commitment*, *reciprocity,* and *liking* might describe subjects that are less susceptible to external manipulation. Instead, they are either motivated by setting and communicating their own goals (*commitment*), by returning a favor they received (*reciprocity*), or complying with recommendations because they like their originators. The latter principles have in common that they describe intrinsic motives to comply with a persuasive attempt, whereas *authority* and *social proof* are related to external pressure and conflict-avoiding motives to comply. This view is further supported by the weak correlation between *social proof* and *commitment*. Persons who are motivated by setting their own goals are evidently not susceptible to social pressure. On the other hand, people who wish to comply with social norms show only a weak tendency to set and pursue their personal goals. *Scarcity* shows only weak correlations with other factors. This is not surprising, since *scarcity* - in contrast to the other principles - does not leverage interpersonal or social mechanisms to motivate a person to comply, but creates a matter-of-fact situation, which is perceived differently by different individuals. This leads to the conclusion that the correlations between factors are supported by theoretical considerations, which confirms the nomological validity of the underlying items. Nomoligical validity "represents the degree to which predictions based on a concept are confirmed within the context of a larger theory" (Bagozzi 1979). It requires that interrelations between constructs are congruent with the underlying theory (Homburg and Giering 1996).

Table 36: Correlation Matrix

	Factor					
	Authority	**Commitment**	**Social Proof**	**Liking**	**Reciprocity**	**Scarcity**
Authority	1	0.103	0.322	0.114	0.257	0.192
Commitment	0.103	1	0.017	0.232	0.376	0.063
Social Proof	0.322	0.017	1	0.205	0.196	0.198
Liking	0.114	0.232	0.205	1	0.222	0.123
Reciprocity	0.257	0.376	0.196	0.222	1	0.172
Scarcity	0.192	0.063	0.198	0.123	0.172	1

To summarize, the results of the Exploratory Factor Analysis are to some degree ambiguous. *Authority*, *commitment*, *social proof*, *liking*, and *reciprocity* fulfill the quality criteria of EFA. In contrast to that, *scarcity* shows relatively low factor loadings and a questionable reliability. However, as this questionnaire is intended for personality profiling and not for measuring the effects under investigation, we conclude that reliability and validity are sufficient for our purpose.

III.4.2 Instrument Development: Persuasive Messages

In the treatment phase of the experiment, SMS messages were sent to the subjects as persuasive interventions. Depending on treatment condition and persuasive profile, the messages were selected such that they represent a fitting, a non-fitting, or a randomly chosen principle. Therefore, a set of messages had to be developed, which implement the different persuasive principles. The set had to be large enough to send different messages to the subjects over time in order to avoid habituation effects. Since a set of context-specific persuasive messages was not available, an expert group on persuasive technologies developed an initial set of 46 messages. The messages were restricted to four out of the six persuasive principles, namely *authority*, *commitment*, *social proof*, and *scarcity*. For the remaining two principles *reciprocity*, and *liking*, the expert group agreed that appropriate messages cannot be formulated. *Reciprocity* would require to do a favor, which makes its receiver feel obliged to return. To implement the principle of *liking*, a personal relationship would be necessary that is perceived positively. A simulation of these principles via text messages seemed to be too equivocal to be incorporated into this study, and therefore, a restriction to four principles was preferred.

To validate that each message implements the intended principle, a web-based card sorting procedure has been applied (Coxon 1999; Harloff 2005). Card sorting has been proposed as an alternative for exploratory factor analysis to assess the validity of measurement scales (Kinicki et al. 1986; Noble and Mokwa 1999; Santos 2006). For this purpose, a panel of experts assigns test items to theoretical constructs. If the items adequately represent the underlying constructs, judges should correctly assign items to the foreseen categories. Consequently, an item is discarded if the number of correct assignments falls short off a pre-defined threshold. The approach has been adopted such that 10 persons not involved in the study had to assign the 46 initial messages to the four persuasive principles or to an "other" category if none of the principles seems to be applicable. In accordance with previous work (Bernardin 1977; Kinicki et al. 1986), the cut-off criterion was defined such that a message is discarded if less than 60% of its assignments fall into the expected category. Consequently, 9 messages were eliminated, which resulted in a set of 37 validated messages. Table 37 shows the original message set and the card sorting results.

Table 37: Persuasive Messages (Validation Results)

Persuasive Message	Authority	Commitment	Social Proof	Scarcity	Other
Governments around the world are fighting overweight as a major cause of disease and premature death. Be part of it and stop snacking.	70	20	10		
Renowned researchers have found an association between snacking between meals and obesity.	80		10		10
Dietitians advise to have 3 meals a day without snacking. Try to reduce snacking.	70	10	20		
Try not to snack today. According to the College of Physicians this is an easy way to lead a healthier life.	100				
Try not to snack today as Dr. Bernstein Diet & Health Clinics advises its customers to avoid snacks during the day to keep a healthy balance.	90				10
Fitness coaches of film stars say that stopping to snack is a prerequisite for a successful fitness plan.	90		10		
The World Health Organization advices not to snack. Snacking is not good for you.	90		10		
In governmental reports experts repeatedly recommend to reduce snacking as a first step to live a healthier life.	90		10		
Following a recent study of the World Health Organization, overweight causes heart diseases, cancer, and diabetes. Do not snack today.	90	10			
A recent study at the University of Amsterdam found out that a weight loss of 3 kg is sufficient for much better well-being. Do not snack.	90	10			
Snacking is the major cause for weight gains, as a recent study of University of Amsterdam found. Eat three times a day and do not snack.	90	10			
According to Weightwatchers snacks can seriously increase obesity.	90	10			
Magazines like Men's Health advice to reduce snacking to have a healthy diet.	80		20		
We all like snacks. And we all know it is the major cause for overweight. So let us stop snacking - even if it's hard. We will be rewarded.		50	40		10
You admire the slim shape of some of your friend? Ask them. They won't take in snacks. Eat three times a day and nothing in between.		40	50		10
It is well-known that the ingredients of most snacks harm your body.	10	10	40		40
Most of the people that successfully lost weight avoided daily snacks. Try to reduce snacking.	10	30	60		
Try to reduce snacking. That is with what most of the participants start to feel more healthy.		30	70		
People that reduce snacking consistently say that they not only feel healthier but also better than before.	10	20	70		
Many people say that they don't spend money for unhealthy nutrition.		20	60	10	10
As much as 47% of employees say that office snack options keep them from eating healthily.		10	70	10	10
90% of people benefits from reducing snacking between meals. It will boost your energy and you will live a healthier life.		10	90		
For a study we asked 50 people about the major cause for their overweight. Overwhelming 80% said it's too much snacks.	10	10	80		

Reduce snacking. You are not on your own: 95% of participants have already reduced snacking.		10	90		
Everybody agrees: not snacking between meals helps you to stay healthy.			90	10	
Already noticed? Most of your friends recognize that snacks are a major cause for overweight. Fight your overweight and stop snacking.			90	10	
Break the vicious circle. Don't let addictive ingredients take control over your life-style.		50			50
How do you feel today? Increase your well-being by a healthy life style. Fresh food, no snacks, and do some sports. What may hinder you?		60		10	30
Did you really reduce snacking the last days? Remember your goal.		80			20
Now it is time to show your commitment in living a healthier life. Avoid snacks to-day.	10	80			10
Try to obtain your goal for living a healthier life by not snacking. You are commit-ted!	10	90			
You have to continue what you've started: you are participating in this test to lead a healthier life. Reduce snacking.	10	90			
By participating in this test you committed yourself to eat healthier. Start already today by not snacking during the day!	10	90			
Remember why you participate in this study. You want to live healthier and lose some weight. The best way is: no snacks.		100			
You can further improve. Try to achieve the goals that you set in the beginning of the study.		100			
The longer you follow your unhealthy habits, the harder it will become to give them up. Don't say 'tomorrow'. Say 'today'.	10	40		40	10
If you feel tempted – resist. It will not be easier the next time.	10	60		20	10
The aim of this study is to live healthier. Reducing snacking is a way to achieve that.		40	10	10	40
Today is a unique opportunity to lead a healthy life. Reduce snacking.		20		40	40
Do not let go away the chance to feel healthy today. Avoid snacks today.		20		50	30
You want to adopt a healthier life-style? Now is the time. Don't wait for tomorrow.		40		60	
You have only a short time in this 3 weeks test to succeed in living healthy. Take your chance today and reduce snacking.		10		90	
There is only one chance a day to reduce snacking. Take that chance today!	10	20		60	10
This test lasts only 3 weeks: you have the unique opportunity to enhance your health by reducing snacking.	10	20		70	
Health is the most precious thing you have. Overweight is the number 1 enemy of health. So don't snack, eat healthy!	10	20	10	60	
The older you get, the harder it becomes to lose weight. It will never again be as easy as today. So no snacks today.		40	10	50	

<u>Notes:</u>

The last five columns contain the % of assignments to each category.

Messages marked grey exceed the threshold of 60% correct assignments.

III.4.3 Sample

III.4.3.1 *Profiling Questionnaire*

Study participants were recruited from a panel by a market research institute in August 2010. Financial compensation was granted for a complete participation in all phases of the experiment. 333 subjects completed the web-based profiling questionnaire. The sample is sufficiently balanced across different age groups, but gender distribution is distorted towards female participants (Table 38).

Table 38: Sociodemographic Sample Structure

Variable	Category	Frequency (N=333)	Percent
Age	18 - 24 years	58	17.4%
	25 - 33 years	81	24.3%
	34 - 42 years	91	27.3%
	43 - 51 years	52	15.6%
	52 - 60 years	32	9.6%
	61 - 73 years	19	5.7%
Gender	female	208	62.5%
	male	125	37.5%

Based on the results of the Principal Component Analysis, factor scores were calculated for each of the six persuasive principles. Factor scores are the scores "that would have been observed for a person if it had been possible to measure the latent factor directly" (Brown 2006). In practical research these scores are often calculated as so-called *coarse factor scores*, i.e. unweighted composites of the raw indicators like sums or averages. However, Monte Carlos studies have shown that such coarse factor scores may be poor representations of the underlying factors (Grice and Harris 1998; Grice 2001). Particularly, factor correlations may be distorted by applying these simple techniques. Therefore, a regression on the correlation matrix from the Principal Component Analysis has been applied to calculate the so-called *refined factor scores*, which usually are less biased than coarse factor scores (Brown 2006; Grice 2001).

Having retrieved the refined factor scores, the persuasive principle with the highest factor score was assumed to be most effective for each individual. As only four out of the six possible principles could be implemented in the experiment, this assignment was restricted to these four principles. Table 39 shows the frequency of the assignments for all six principles and the restricted set of four principles. Evidently, none of the principles is preferred in the sample population as the assignments are quite balanced across the six as well as across the four principles.

Table 39: Frequency of Assignments to Persuasive Principles

Principle	Six Principles		Four Principles	
	Frequency	Percent	Frequency	Percent
Authority	58	17.4 %	75	22.5 %
Commitment	52	15.6 %	83	24.9 %
Social Proof	66	19.8 %	92	27.6 %
Scarcity	57	17.1 %	83	24.9 %
Liking	51	15.3 %		
Reciprocity	49	14.7 %		

III.4.3.2 ***Experiment***

After having completed the profiling questionnaire, all 333 respondents were invited to participate in the persuasive experiment. Out of these 333 respondents, 112 started with the experiment and entered at least one entry in the online diary. Before the start of the experiment, each respondent was assigned to one of the three experimental conditions (i.e. WRONG, RIGHT, or RANDOM). From the 112 subjects who began with the experiment, 33 had been assigned to the WRONG condition, 42 to the RIGHT condition, and 37 to the RANDOM condition.

Persuasive treatments usually require several inventions until a behavioral change occurs (Prochaska and DiClemente 1992). We therefore eliminated all subjects from the data set that had not entered at least 2 measures in each of the two experimental phases (baseline and treatment phase). This resulted in a final data set of 476 repeated measures from 55 subjects. Table 40 shows the number of subjects and measurements under each condition in the two phases. Although the number of subjects and measures are similar across the different conditions and phases, the differences are too large to be considered as a balanced sample. Furthermore, the number of measures within each subject and phase varies between two and seven. Both aspects have to be taken into account for selecting an adequate evaluation method because balanced samples and a fixed number of repeated measures is a prerequisite for the applicability of a number of multivariate analysis techniques (see section III.5.1).

Table 40: Number of Subjects and Measures in the Experiment

Condition		Baseline Phase	Treatment Phase	Sum
WRONG	# Subjects	16	16	(*)
	# Measures	64	68	132
RIGHT	# Subjects	21	21	(*)
	# Measures	87	99	186
RANDOM	# Subjects	18	18	(*)
	# Measures	81	77	158
Sum	# Subjects	55	55	(*)
	# Measures	232	244	476
(*) Sum not applicable because subjects are the same in both phases				

III.5 Data Analysis

III.5.1 Analysis Methodology

III.5.1.1 *Method Selection*

Longitudinal data, as obtained from the experiment in the present study, are often analyzed by basic multivariate regression and analysis of variance methods such as repeated-measures ANOVA. However, these techniques impose assumptions on the experimental data set, which are often violated in research practice.

Most multivariate techniques require balanced samples, which means that an equal number of measures are taken for all subjects and that the number of subjects is equal across all experimental conditions. Otherwise the application of these techniques will distort results in an unpredictable way, particularly if the data is not normally distributed (Wilcox 2005). In addition to this, basic multivariate techniques impose the assumption of *sphericity*, which means that the "variances of the differences between treatment levels" must be equal (Field 2009).

As demonstrated in section III.4.3.2, our sample is unbalanced with regard to the number of subjects per condition and the number of measures taken per subject. To obtain a balanced sample, we could omit subjects such that the number of subjects is equal across the three experimental conditions. To overcome the problem of unbalanced measures, averages can be calculated for each condition and phase. Applying this approach, a highly significant Mauchly's test for sphericity ($p<0.001$) has shown that multivariate techniques are not appropriate for the present data set. In addition, such simplistic techniques severely reduce analysis quality because they neglect the complexity inherent in longitudinal data, increase standard errors, and may cause bias by omitting subjects and measures (for a detailed discussion of problems arising from such data reduction techniques, see Weiss (2005)).

Accurate analysis techniques must be able to deal with the typical characteristics inherent to longitudinal data, which are discussed in the following (Landwehr et al. 2008).

Between-subject variability

Repeated measures taken over time will usually differ between subjects. In the present study, we can expect that subjects differ in their general nutrition behavior. Some subjects will regularly take in many snacks each day, whereas for others snacking will be a rare exception. Additionally, the number of snacks, which is the measured variable, is not clearly defined. Eating five biscuits may be regarded by one subject as a single snack, whereas another subject may report it as five snacks. Therefore systematic between-subject variability can be expected in the present data. This will lead to a high error variance, which may result in an underestimation of the treatment effects (Gueorguieva and Krystal 2004; Quené and van den Bergh 2004).

Correlated measurement errors

Between-subject variability is a consequence of unobserved or unobservable factors and is therefore captured as measurement errors. As the unobservable factors, which are responsible for between-subject variability, are mostly stable over time (because they are deeply rooted for example in personality traits, cognitive predispositions or behavioral patterns), measurement errors belonging to the same subject will be correlated more highly than measurement errors belonging to different subjects. Neglecting the correlation between error terms results in an underestimation of standard errors and therefore to an overestimation of statistical significance (Gujarati 2003).

Autoregressive correlation

In longitudinal data sets, adjacent measures within subjects are often more highly correlated with each other than more distant observations because the change of latent variables is not constant over time (Fitzmaurice et al. 2004). This so-called *autoregressive correlation* further differentiates correlated error terms. If a statistical model does not consider autoregressive correlations but assumes constant correlations between error terms, standard errors are underestimated, which leads to overestimated statistical significance (Gueorguieva and Krystal 2004).

Heteroscedasticity

Heteroscedasticity is present when the residuals at each level of the predictor variables have unequal variances (Field 2009). Varying variances across different experimental groups may occur because a treatment often reduces the variance compared to non-treatment (Fitzmaurice et al. 2004). In our sample, a "RIGHT" treatment might theoretically prevent subjects completely from snacking, whereas "WRONG" treatment might have no effect. Before the treatment, each experimental group should have a similar variance. After treatment, the WRONG group would still have the same variance, whereas the RIGHT group would have a variance of zero. This hypothetical scenario illustrates that heteroscedasticity is likely to be present in our sample. If heteroscedasticity is ignored, standard errors are overestimated, which results in underestimated statistical significances (Gujarati 2003).

A number of approaches, often referred to as *panel analysis methods* (Landwehr et al. 2008; Schröder 2006), have been proposed to extend linear models to address certain characteristics of longitudinal data:

- *Pooled Regression* still assumes that error terms are independent and that heteroscedasticity is not present, which are unrealistic assumptions for experimental repeated measures (Gujarati 2003).
- *Dummy Variable Regression* accounts for between-subject variances but delivers inconsistent estimations if only few measures are taken for each subject. Furthermore, autoregressive correlations are not considered (Schröder 2006).
- *Random-Effects Models* incorporate an accurate treatment of between-subject variability and correlated error terms, but do not account for autoregressive correlations and heteroscedasticity (Landwehr et al. 2008).

To summarize, many statistical evaluation methods, which are usually applied for analyzing longitudinal data, are not appropriate for the given data sample. Gueorguieva and Krystal (2004) and Landwehr et al. (2008) provide a more detailed discussion of different methods.

As an alternative, *Linear Mixed Models* (LMM) are an appropriate analysis method to overcome the diverse issues associated with the characteristics of longitudinal data. They do not require balanced samples or an equal number of measures for all subjects (Verbeke and Molenberghs 2000). Sphericity, homogeneity of regression slopes, or independence of observations are not assumed (Field 2009). Furthermore, LMMs account for between-subject variability, correlated error terms, autoregressive correlations, and heteroscedasticity (Fitzmaurice et al. 2004; Gueorguieva and Krystal 2004;

Verbeke and Molenberghs 2000). Finally, independent variables in LMMs may involve a mix of fixed and random effects, which are estimated on the basis of a linear model (this characteristic led to the name *Linear Mixed Models*).

The given sample can be expected to violate a number of assumptions of classical multivariate techniques. Therefore LMMs have been chosen as statistical evaluation method. LMMs are implemented in a number of software packages (e.g. SPSS, SAS, R, STATA). For the present study, the MIXED function of SPSS was used. In the following section we explain the fundamentals of LMMs, followed by a section on their practical application.

III.5.1.2 ***Fundamentals of Linear Mixed Models***

Linear Mixed Models are parametric linear models for clustered or longitudinal data (West et al. 2007). They estimate the relationship between a continuous dependent variable and several continuous or categorical predictor variables. LMMs can include *fixed effects* and *random effects*. Fixed effects describe the relationships of predictor variables on dependent variables for the whole population (population effects). Typical examples for fixed effects are treatments or experimental conditions. They have to include all levels that are of interest for the investigation.

Random effects describe relationships that are specific to clusters or subjects within the population. They model the random deviations from the fixed-effect relationships between different levels of the data (West et al. 2007), which reduces the standard error of parameters and consequently the β-error of the estimation (Gueorguieva and Krystal 2004).

The LMM approach can be described as a sequence of two analysis steps (Fitzmaurice et al. 2004; Verbeke and Molenberghs 2000). In a first step, regression coefficients are calculated for the repeated measures of each individual subject. The resulting matrix of regression vectors represents a set of subject-specific regression lines for the dependent variable over time. In a second step, the set of individual regression vectors is condensed to a "mean" regression vector representing a mean regression line for the whole sample. Different fixed effects can be differentiated by combining the individual regression vectors along different experimental conditions (Verbeke and Molenberghs 2000).

In matrix notation, an LMM for an individual subject is specified by the following equation (West et al. 2007):

$$\mathbf{Y}_i = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{u}_i + \boldsymbol{\varepsilon}_i$$

In this equation, $\mathbf{Y}_i$ is the vector of longitudinal observations on the dependent variable for subject i. The number of elements in vector $\mathbf{Y}_i$ can vary from subject to subject. $\mathbf{X}_i$ is a matrix of the known covariates associated with the unknown fixed-effects coefficients represented by vector $\boldsymbol{\beta}$. The number of rows in $\mathbf{X}_i$ is equivalent to the number of observations, and the number of columns is equivalent to the number of covariates. As fixed effects relate to the whole sample population, $\boldsymbol{\beta}$ does not contain a subject-specific index. The matrix $\mathbf{Z}_i$ represents the known covariates associated with unknown random-effects coefficients represented by vector $\mathbf{u}_i$. As random effects are subject specific, this vector contains the subject-specific index i. Vector $\boldsymbol{\varepsilon}_i$ contains the residuals associated with each observation on the dependent variable for subject i.

Random effects $\mathbf{u}_i$ and residuals $\boldsymbol{\varepsilon}_i$ are defined to be random variables with the following assumptions:

$$\mathbf{u}_i \sim N(\mathbf{0}, \mathbf{D})$$

$$\boldsymbol{\varepsilon}_i \sim N(\mathbf{0}, \mathbf{R}_i)$$

This means that the random effects $\mathbf{u}_i$ follow a multivariate normal distribution with a mean vector of $\mathbf{0}$ and a variance-covariance matrix of $\mathbf{D}$, and the residuals $\boldsymbol{\varepsilon}_i$ follow a multivariate normal distribution with a mean vector of $\mathbf{0}$ and a variance-covariance matrix of $\mathbf{R}_i$.

The assumption that the random effects have a mean value of 0 is equivalent to the interpretation that random effects represent the subject-specific deviation from the marginal population mean (Landwehr et al. 2008). Unsystematic deviation due to between-subject variability is captured by the $\mathbf{D}$ matrix, which describes the random effects in the model, whereas the $\mathbf{R}_i$ matrix represents deviation attributable to within-subject variability. An appropriate selection of a covariance structure for $\mathbf{R}_i$ allows for an accurate adaptation to the observed data. As described in section III.5.1.1, longitudinal data often show autoregressive correlations. This characteristic can be accounted for by specifying an autoregressive covariance structure (other covariance structures are described, for example, in West et al. (2007)).

Once the model is specified, model parameters are estimated in a two-step procedure. First, the parameters for the $\mathbf{D}$ and $\mathbf{R}_i$ matrices are estimated by means of Maximum-Likelihood methods. Then fixed effects are estimated by applying the parameter estimations from the first step for weighting a Generalized Least Squares model (West et al. 2007). Applying the standard Maximum Likelihood (ML) algorithm to estimate the parameters for **D** and $\mathbf{R}_i$ results in a bias because ML does not account for the loss of degrees of freedom arising from the subsequent estimation of fixed effects (Verbeke and Molenberghs 2000). Therefore an alternative algorithm called *REML* (*Restricted Maximum Likelihood* or also *Residual Maximum Likelihood*) is usually recommended, which produces unbiased parameter estimates by correcting degrees of freedom by the number of subsequently estimated fixed effects (West et al. 2007).

III.5.1.3 *Practical Application of LMMs*

LMMs offer a plethora of options to adapt a model to a given data set. As explained before, the investigator has to make choices about which fixed and random effects to include in the model, whether to relate random effects only to the slopes of a regression or also to the intercepts, and about which covariance structures to apply to the random effects. To find the "best" model among the many competing options, different pairs of models are iteratively estimated and compared to each other. In each iteration, a statistical test is applied to decide, which of the two tested models is preferred with regard to model fit and parsimony. The goal of this iterative process is to find the model that best predicts the dependent variable and at the same time is parsimonious with regard to the number of parameters used (West et al. 2007).

Different statistical tests and model-fit criteria are available (West et al. 2007). Among them, the so-called *Likelihood Ratio Test* (LRT) holds an outstanding position because it implies a statistical test of whether one model is significantly better than another one. For comparing two models with LRT, one model must be *nested* within the other one, which means that one model is a generalization of the other one. The nested model then is a special case of the more general model. A nested model can be obtained by imposing constraints on the parameters in the nesting (or *reference)* model. For a given model with fixed and random effects, a nested model may be obtained by omitting a random effect, which is equivalent to the constraint of setting the associated parameters to zero. To conduct an LRT between a nested and an associated reference model, the following test statistic is calculated (Field 2009; West et al. 2007):

$$\chi^2_{\text{Change}} = -2 \log \left(\frac{L_{nested}}{L_{reference}} \right) = -2 \log (L_{\text{nested}}) - (-2 \log (L_{\text{reference}})) \sim \chi^2_{\text{df}}$$

where df = (Number of Parameters)$_{\text{reference}}$ - (Number of Parameters)$_{\text{nested}}$

L_{nested} and $L_{reference}$ are the values of the likelihood function produced by the ML or REML algorithm. By default, they are provided by software packages like SPSS. The test statistic χ^2_{Change} is asymptotically χ^2-distributed with *df* degrees of freedom, which is obtained by subtracting the number of parameters in the nested model from the number of parameters in the reference model (the number of parameters is also provided by SPSS and other software packages). In general, smaller values for -2 log L_x indicate better model fit. To test whether the difference between two models is statistically significant, a χ^2-test is applied on χ^2_{Change} with *df* degrees of freedom.

For models that differ in their covariance parameters (i.e. they have different random effects or covariance structures), LRT can be applied to ML and REML estimations. As REML produces more accurate estimates, REML should be used in this case. However, if the models differ in their fixed-effect parameters, REML must not be used. Instead, ML estimation is mandatory to compare the models with an LRT.

To test models that differ in their covariance parameters, two cases have to be distinguished. If the covariance parameters differ in the presence or absence of a random effect (case 1), the associated variance parameter of the nested parameter is at the boundary of the parameter space. Omitting a random parameter is equivalent to constraining its variance to zero. Since variances cannot be smaller than zero, this is the boundary of the parameter space. Therefore the χ^2-test must be one-tailed, which means in practice that the significance level must be divided by two (Landwehr et al. 2008; West et al. 2007). If the covariance parameters vary in terms of different covariance structures (i.e. different covariance matrices R_i) specified for the two models (case 2), a (two-tailed) LRT (preferably on the basis of REML estimates) is valid under the condition that the two covariance structures have a nesting relationship.

To compare models that do not have a nesting relationship, LRT must not be applied, and alternative procedures that allow for statistical inference are not available. Instead, model fit may be assessed by comparing information criteria calculated for the different models. Two information criteria are usually applied, namely the *Akaike Information Criterion* (AIC) and the *Bayes Information Criterion* (BIC), which are both calculated by LMM software packages. As none of the two criteria has turned out to be superior with respect to each other (Gurka 2006), both should be considered in model selection.

In research practice, finding the best models usually requires several iterations, in which an appropriate mix of theoretical considerations and statistical procedures has to be applied. The purpose of these iterations is to identify the best combination of fixed effects, random effects and covariance structures while keeping the model as parsimonious as possible. Due to the diversity of potential applications, there is no single model-building strategy that fits to all situations. However, two basic strategies have been described in literature, which may also be wisely combined in practice (Verbeke and Molenberghs 2000; West et al. 2007). The first strategy, often referred to as *top-down strategy*, starts with a model that includes the maximum number of potential fixed effects. The second strategy, often called *step-up strategy*, starts with only one intercept as fixed parameter and several higher-level random effects. The top-down approach has been found to be more efficient for longitudinal data and is therefore applied in the present study. It consists of the following four basic steps (Landwehr et al. 2008; Verbeke and Molenberghs 2000; West et al. 2007):

Step 1: Specification of fixed effects

On the basis of theoretical considerations, all relevant fixed effects and interactions are added to the model.

Step 2: Specification of random effects

Random effects are iteratively added to the model. In each iteration, only one effect is added. A one-tailed LRT is applied to decide whether the extended model fits the empirical data significantly better than the more parsimonious model without the newly added random effect. Furthermore, different covariance structures for the D-matrix can be tested in this step.

Step 3: Selection of an appropriate covariance structure for the residuals

On the basis of the best-fitting model identified in step 2, an appropriate covariance structure for the residuals (R-matrix) is iteratively determined. In case of nested models, improvements are assessed by a two-tailed LRT. In case of non-nested models, the information criteria AIC and BIC have to be applied. For longitudinal data, the appropriateness of an autoregressive covariance structure should be tested in this step.

Step 4: Model reduction

In the final model from step 3, non-significant fixed effects are assessed with respect to their contribution to the model fit. The significance of fixed effects is usually tested by Type III F-Tests, the results of which are included in the result report of SPSS and other statistical software (alternatives are Type I F-Test, z-Test, t-test, Wald χ^2-Test; see West et al. (2007) for a discussion of these tests). Iteratively, fixed effects are omitted from the model. An LRT on the basis of ML estimates reveals whether the model should further include a fixed-effect although it is not significant. REML estimation must not be used in this analysis step.

Once the optimal model is found, REML estimation is conducted to obtain the final results for further interpretation. Similarly to linear regression techniques, fixed effect coefficients are typically interpreted to analyze the main experimental effects. Random effects allow for conclusions such as how certain random conditions could affect the experimental effects.

III.5.1.4 *Implementation of the Top-Down Model Building Strategy*

To find the "best" LMM for our experimental data, a top-down fitting procedure has been applied as illustrated in Figure 10.

Step 1: Specification of fixed effects

To implement the top-down procedure for the given sample, we first fit a model, which only includes the fixed effects for CONDITION, PHASE and the CONDITION x PHASE interaction. CONDITION is an indicator variable that describes the experimental condition (0: "WRONG"; 1: "RIGHT"; 2: "RANDOM"). PHASE indicates whether a subject has undergone a treatment (0: no treatment; 1: treatment in accordance with condition). The interaction between the two indicator variables is included to test whether the experimental condition does influence the outcome of the treatment. The corresponding linear model can be formulated as follows:

M1: $SNACKS_{ti} = \beta_0 + \beta_1 \text{ x } CONDITION_{ti} + \beta_2 \text{ x } PHASE_{ti} + \beta_3 \text{ x } CONDITION_{ti} \text{ x } PHASE_{ti} + \varepsilon_i$

$SNACKS_{ti}$ denotes the observed number of snacks for subject i at time t. β_0 is the intercept of the linear model, $\beta_{1,\,2,\,3}$ are the regression coefficients, which have to be estimated, and ε_i denotes the residual for subject i.

Step 2: Specification of Random Effects

In step 2, two random effects are iteratively added and tested. First, a subject-specific random effect is added for the regression intercept (model M2.1), and second, random intercepts for CONDITION are included (M2.2). The hypotheses H_M1 and H_M2 are tested by one-tailed LRTs with REML estimation to decide which of the three models is best-fitting. The two models are specified as follows:

M2.1: $SNACKS_{ti} = \beta_0 + \beta_1 \text{ x } CONDITION_{ti} + \beta_2 \text{ x } PHASE_{ti} + \beta_3 \text{ x } CONDITION_{ti} \text{ x } PHASE_{ti} + u_i + \varepsilon_{ti}$

M2.2: $SNACKS_{ti} = \beta_0 + \beta_1 \times CONDITION_{ti} + \beta_2 \times PHASE_{ti} + \beta_3 \times CONDITION_{ti} \times PHASE_{ti} + u_{0i} + u_{1i} \times CONDITION + \varepsilon_{ti}$

The term u_{0i} represents the random intercept associated with subject i. u_{1i} denotes the random effect for CONDITION within each subject. For M2.1, a *Variance Component* (VC) matrix is specified as there is only one random effect, and therefore covariances cannot occur. In M2.2, an *unstructured* matrix is applied to allow for any variances and covariances.

Step 3: Selection of an appropriate Covariance Structure for the residuals

A commonly used covariance structure for longitudinal data is the *first-order autoregressive* structure (AR1). This structure implies that adjacent data points are more correlated than data points which are further apart from each other. At the same time, AR1 assumes homogeneous variances across observations (West et al. 2007). Model M3.1 relaxes the model M2.1 such that an AR1 covariance structure is applied to the residuals.

Model M3.2 further relaxes model M3.1 by allowing for heterogeneous variances across observations. For this purpose, a heterogeneous *first-order autoregressive* structure (AR1: Heterogeneous or ARH1) is applied.

To test the superiority of the ARH1 structure over a more parsimonious structure that assumes uncorrelated observations within each subject, a *diagonal* structure for the **R** matrix is applied in model M3.3. The diagonal matrix allows for differing variances between observations but fixes the covariances between observations to zero. The corresponding hypotheses H_M3, H_M4 and H_M5 are tested by two-tailed LRTs with REML estimation. Practically, the different covariance structures are declared by adding a */REPEATED* statement to the *MIXED* command in SPSS (West et al. 2007).

Step 4: Model Reduction

As will be shown in the next section, CONDITION is the only effect that is not significant. Since eliminating CONDITION alone would neither change the number of parameters in the model nor the -2Log(L) value, the fixed effects for CONDITION and the PHASE x CONDITION interaction are omitted in this step. Consequently, we test a model in which only the experimental phase influences the outcome variable versus a model in which the outcome depends on the experimental phase, the condition, and the interaction of these two independent variables. As we are comparing models with different fixed effects in this step, ML has to be applied instead of REML estimation. Hypothesis H_M6 is tested by a two-tailed LRT.

Once the best-fitting model is found, a final REML estimation is applied to obtain parameter values for further interpretation (see section III.5.4).

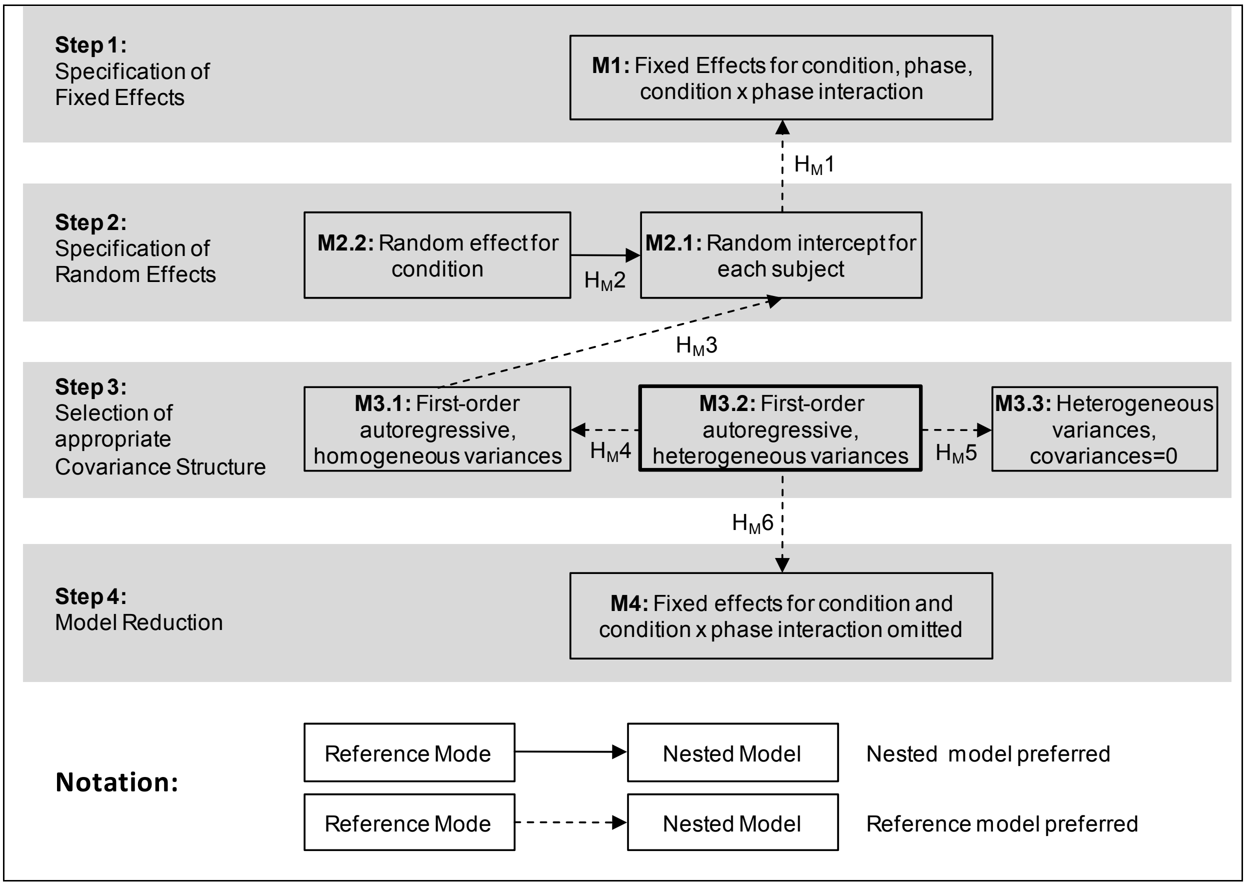


Figure 10: Top-Down LMM Fitting Procedure

III.5.2 Model Fitting

Following the model fitting procedure described in section III.5.1.4, we first estimated a model that only contained the fixed effects of CONDITION, PHASE, and their interaction (model M1).

In step 2, this model was compared to a model which contains random intercepts for each subject. The corresponding hypothesis H_M1 could be confirmed by an LRT, which means that M2.1 better explains the present sample. We can conclude from this confirmation that the intercepts vary significantly across subjects. Hypothesis H_M2 could not be confirmed; therefore a random effect for CONDITION is not included in the final model. This could be expected because a random assignment to the three conditions has been applied; so there should not be a systematic variance across the subjects in the three conditions.

In step 3, we applied three different covariance structures to the residuals of M2.1. A first-order autoregressive covariance structure with heterogeneous variances turned out to best fit the underlying sample data (H_M3 and H_M4 were confirmed, H_M5 was not confirmed).

Among the three fixed effects, CONDITION was not significant ($p = 0.175$), whereas PHASE ($p<0.001$) and PHASE x CONDITION ($p = 0.12$) were both significant. In step 4, we therefore omitted the fixed effect of CONDITION. Furthermore, we removed the interaction term from the model because otherwise the number of parameters is not reduced and the -2LL value remains equal (i.e. model fit does not change). An ML-based LRT has shown that this modification reduced the quality of the model. This means that a model, which has PHASE (i.e. treatment vs. no treatment) as the only fixed effect, explains the sample data significantly worse than a model that also takes into account the experimental condition.

To summarize, we conclude from the model fitting procedure that model M3.2 is the best-fitting model among the tested alternatives. The model fitting results for steps two to four are summarized in Table 41.

Table 41: Model Fitting Results

Step	Hypothesis	-2LL (nested)	df (nested)	-2LL (reference)	df (reference)	p	Estimation Method	Superior Model
2	H_M1	1549.304	7	*1449.030*	*8*	<0.001	REML	M2.1
	H_M2	*1449.030*	*8*	1446.548	17	0.491	REML	M2.1
3	H_M3	1449.030	8	*1435.850*	*9*	<0.001	REML	M3.1
	H_M4	1435.850	9	*1383.811*	*23*	<0.001	REML	M3.2
	H_M5	1396.217	22	*1383.811*	*23*	<0.001	REML	M3.2
4	H_M6	1387.688	19	*1373.111*	*23*	0.006	ML	M3.2
Values for the superior model in each iteration are marked in italics.								

III.5.3 Descriptive Analysis

Before the best-fitting LMM identified in the previous section is applied to the data, we explore the results of our experiment by evaluating graphical illustrations and descriptive statistics.

In Figure 11, the mean values of the number of unhealthy snacks are plotted against the experimental phase (i.e. baseline phase vs. treatment phase). The separate lines represent the different experimental conditions (i.e. WRONG, RIGHT, and RANDOM). Inspecting this graph, we can expect from the data set that a strong negative effect of the treatment on snacking behavior will be observed for the RIGHT and RANDOM group. In contrast, even a slightly positive effect is observed for the WRONG group. Against initial expectations, the RIGHT group has reduced their snacks to a lesser degree than the RANDOM group, even if the difference between these groups is small. Both groups had an average of 1.2 snacks per day in the baseline phase. In the treatment phase, the RIGHT group reduced their consumption to 0.68 snacks, which means a reduction by 43%. The RANDOM group achieved a reduction by 55% to 0.54 snacks per day. The WRONG group even increased their number of snacks from 1.27 to 1.4 snacks per day, which means an increase of 10%.

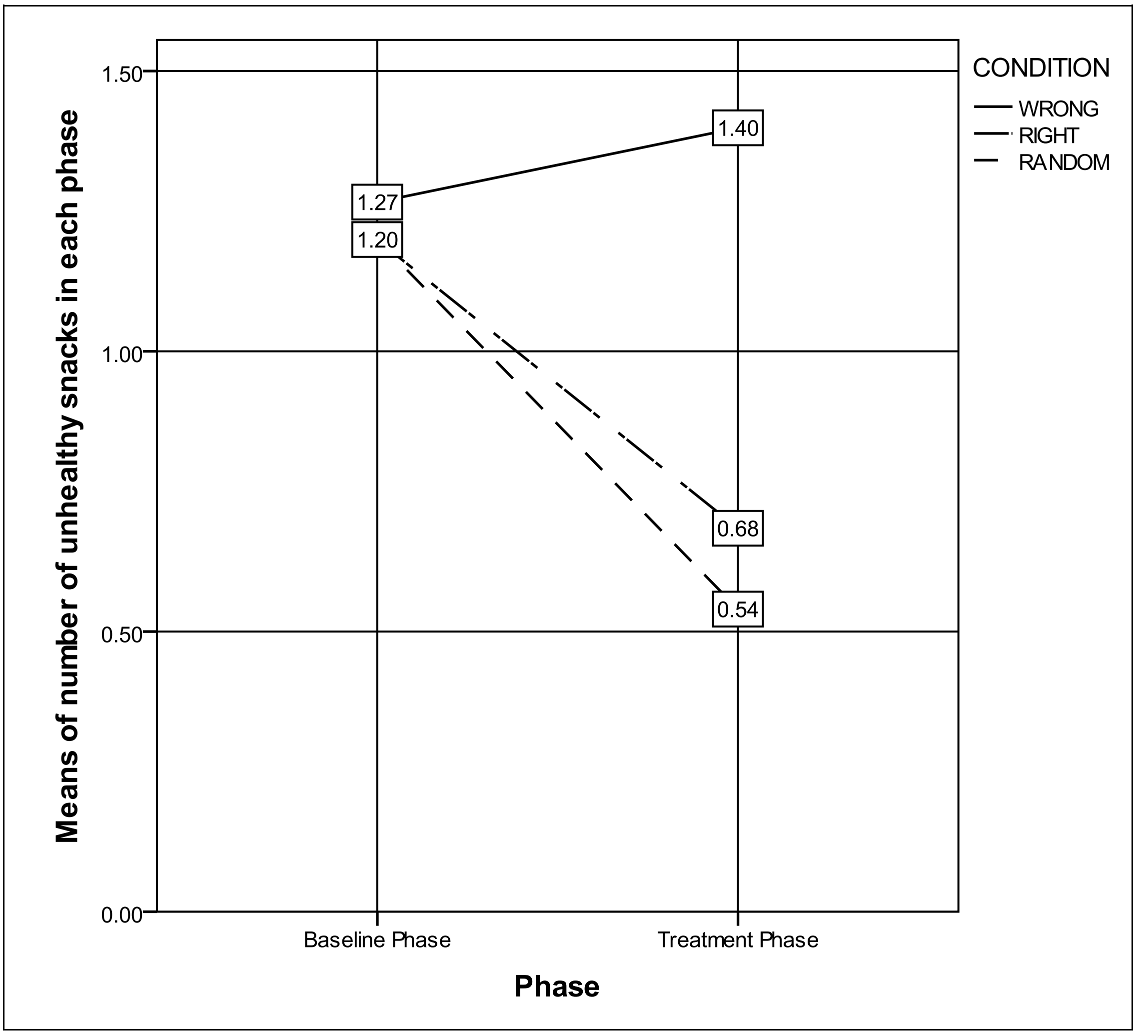


Figure 11: Number of Unhealthy Snacks in each Phase

Figure 12 shows the development of the snacking behavior over time for each experimental group. The time axis in this graph represents the days since the beginning of the experiment. In time point 0, all groups started at almost the same number of snacks per day. Despite some variation, on average the number of snacks remained quite stable in the baseline phase (i.e. neither a positive nor a negative trend can be observed in the three groups).

Starting with day six (i.e. the seventh day of the experiment), subjects received a persuasive message each day. We can see that the number of snacks has decreased sharply in all three groups on this day. This negative trend was continued for the RIGHT and RANDOM group over the following days, but came to a standstill for the WRONG group (with a strong outlier on day 12). Whereas the number of snacks was clearly below the starting point of the RIGHT and RANDOM group, the WRONG group ended the experiment on almost exactly the same level as they started. Furthermore, this graph shows that the RIGHT and the RANDOM group behaved very similar over time, so that a significant difference cannot be expected.

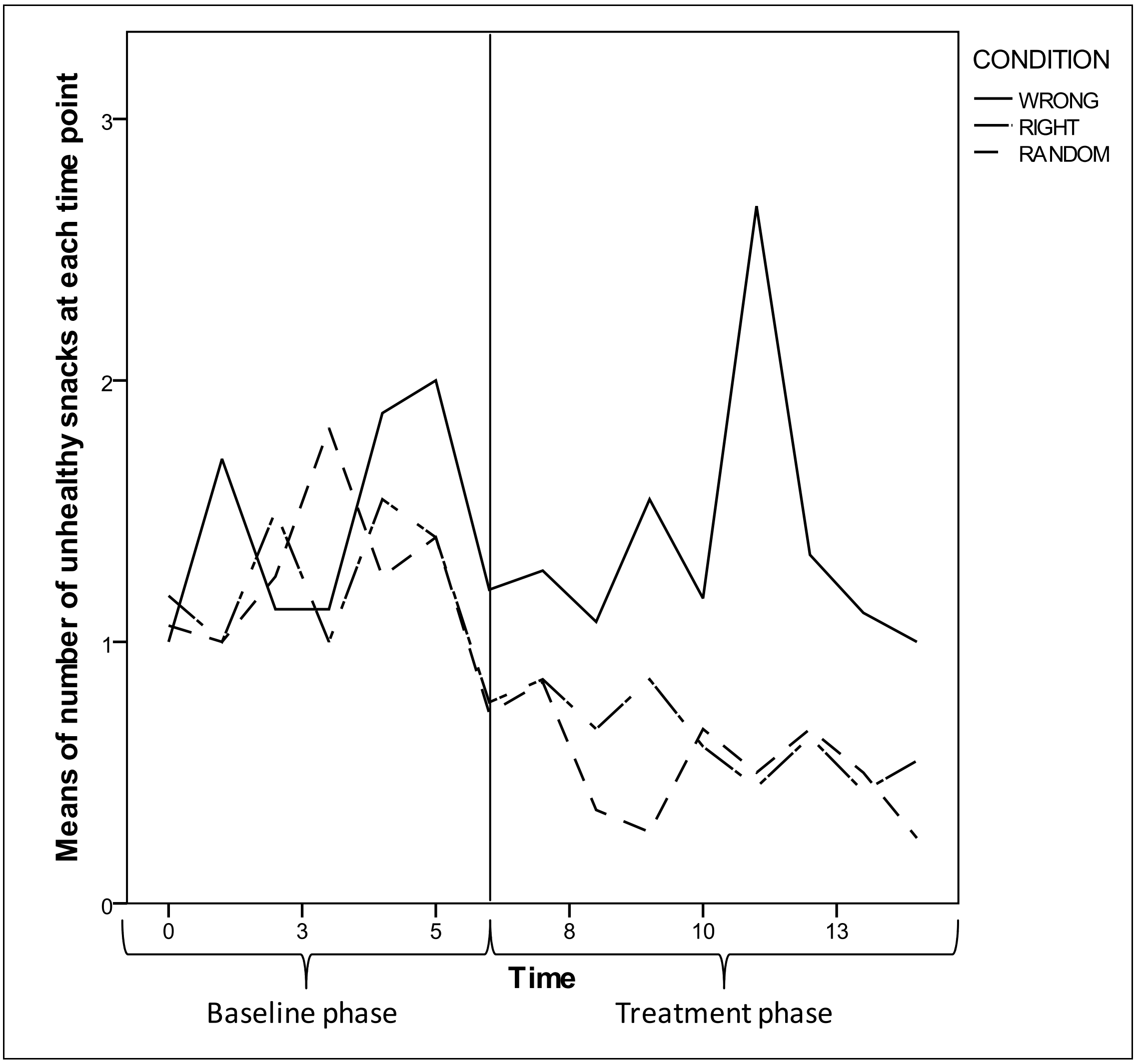


Figure 12: Number of Unhealthy Snacks at each Time Point

Figure 13 shows a scatterplot of the mean number of snacks at each time point and regression lines for the three experimental groups over time. The regression lines corroborate the impression we gained from the previous graphs. Whereas the WRONG group shows an almost horizontal line, the RIGHT and RANDOM groups have a clearly negative trend over time in their snacking behavior. Inspecting the coefficients of determination (R^2) for the regression lines, an R^2 value of close to zero indicates that TIME does not account for the variability in the number of snacks in the WRONG group. As TIME can be interpreted to some degree as an indicator for the strength of the treatment, the extremely low R^2 value indicates that the experimental treatment has a very limited effect on the snacking behavior. At the same time, regression lines for the RIGHT and RANDOM group have R^2 values of 0.633 and 0.558, which means that TIME accounts for 63% respectively 55% of the variability in snacking behavior. This further fosters our interpretation that experimental treatment had a strong effect in the RIGHT and RANDOM group but almost no effect in the WRONG group.

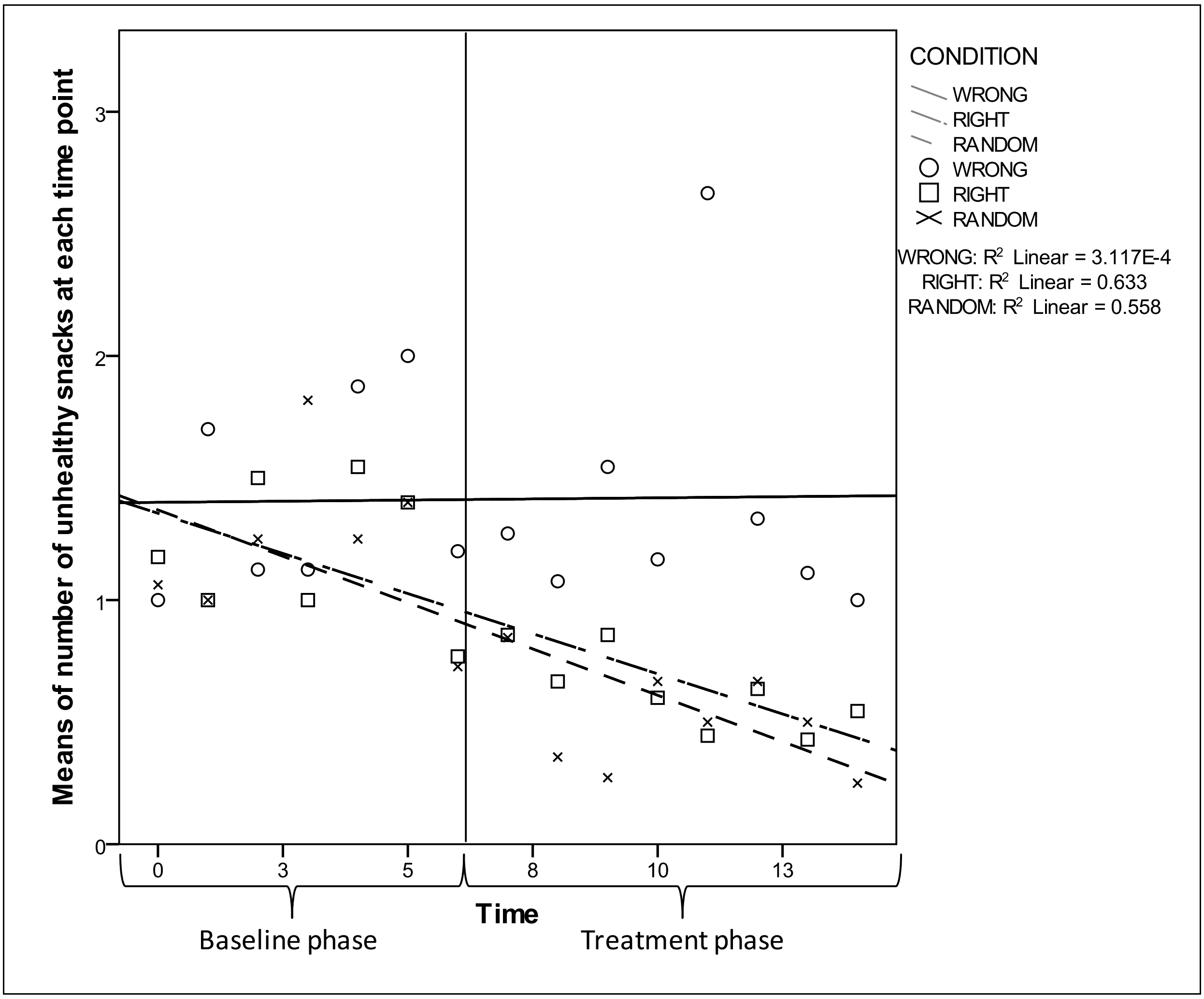


Figure 13: Scatterplot and Regression Lines

To summarize, the graphical evaluations of our data gives strong support for our expectation that persuasive treatment is less effective if an inappropriate persuasive strategy is applied than if an appropriate strategy is adopted. Furthermore, there is strong evidence that there is only a marginal difference between selecting the most appropriate strategy and a random selection of persuasive messages. Finally, descriptive analysis indicates that adopting the most inappropriate strategy might even have a slightly counterproductive effect. In the following section, the assumptions gained from this graphical analysis will be tested by applying the LMM developed in the previous section.

III.5.4 Experiment Evaluation

After having identified the best-fitting model (section III.5.1.4), a final REML estimation is conducted on model M3.2. On the basis of the estimation results, the hypotheses formulated in section III.3.2 are tested. Table 42 presents the results of the significance tests (Type III F-tests) for the fixed effects of CONDITION, PHASE, and CONDITION x PHASE. We can see that CONDITION alone is not a significant predictor for the outcome variable SNACKS. This could be expected as the experimental condition does not account for treatment. Instead, the effect of PHASE is highly significant, which means that the outcome variable significantly varies between the baseline phase and the treatment phase. Furthermore, the significant interaction term shows that the treatment is significantly different for the different experimental groups.

Table 42: Significance Tests of Fixed Effects

Source	Numerator df	Denominator df	F	Sig.
Intercept	1	51.376	102.651	.000
CONDITION	2	51.309	1.804	.175
PHASE	1	166.874	13.182	.000
CONDITION x PHASE	2	166.037	4.584	.012

With these results however, we can neither infer about the direction of the effects nor about the differences between the experimental conditions. Therefore, we next inspect the parameter estimates for the fixed effects (Table 43). As all fixed effect factors are categorical variables, the parameter estimates are represented as contrasts against the reference category, which is RANDOM for CONDITION and TREATMENT for PHASE. Therefore, these parameter levels are set to zero in the contrasts. The contrasts show that the effect of CONDITION alone is significant if we compare WRONG vs. RANDOM, but is not significant if we compare RIGHT vs. RANDOM. However, this finding is not sufficient for the hypotheses to be tested as it does not yet account for the experimental phase.

Next, we see that PHASE has a significant effect on the dependent variable SNACKS. Across all conditions, a change from phase 1 to phase 2 resulted in a reduction of 0.67 snacks per day. Again, we cannot infer from this result to our hypotheses as it does not account for the different experimental conditions. Therefore we next inspect the results for the CONDITION x PHASE interaction.

The contrasts for the different levels of CONDITION show that under the WRONG condition, the effect of PHASE is 0.731 units weaker than under the RANDOM condition. As this effect is significant, our hypothesis **H2** is confirmed. Furthermore, the effect of PHASE under the RIGHT condition is 0.228 units weaker than under the RANDOM condition, but this interaction effect is not significant. Therefore, we reject hypothesis **H3**, which assumes that the RIGHT condition has a positive influence on the effect of PHASE. Instead, we cannot observe a significant difference between the RIGHT and RANDOM condition.

Table 43: Parameter Estimates for Fixed Effects

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						lower bound	upper bound
Intercept	.470015	.174075	57.545	2.700	.009	.121508	.818522
[condition=1]	.799186	.254795	58.328	3.137	.003	.289221	1.309152
[condition=2]	.184098	.235917	56.866	.780	.438	-.288342	.656539
[condition=3]	0	0	.	.	.	.	.
[phase=1]	.671197	.165197	157.774	4.063	.000	.344913	.997480
[phase=2]	0	0	.	.	.	.	.
[condition=1] * [phase=1]	-.731320	.245238	165.584	-2.982	.003	-1.215516	-.247123
[condition=1] * [phase=2]	0	0	.	.	.	.	.
[condition=2] * [phase=1]	-.228387	.226725	164.268	-1.007	.315	-.676058	.219284
[condition=2] * [phase=2]	0	0	.	.	.	.	.
[condition=3] * [phase=1]	0	0	.	.	.	.	.
[condition=3] * [phase=2]	0	0	.	.	.	.	.
Variable encodings: CONDITION = 1 : WRONG; 2 : RIGHT ; 3 : RANDOM PHASE = 1 : BASELINE ; 2 : TREATMENT							

To obtain a contrast between the conditions WRONG vs. RIGHT, the LMM was estimated for a subset of the original data, in which the RANDOM condition was omitted (Table 44). A significant parameter estimate of -0.521 supports our expectation that under the WRONG condition, the treatment effect is lower than under the RIGHT condition. We therefore accept hypothesis **H4**.

Having compared the different conditions against each other, we next assess whether the treatment with persuasive messages has a significant effect on SNACKS under each condition. For this purpose, we conducted a post-hoc test with LSD adjustment (a more conservative adjustment than LSD is not necessary because the PHASE variable has only two levels). This test estimates the so-called *estimated marginal means* (also known as *adjusted group means* or *predicted means*), which are group means estimated from the fitted model. The test then conducts a pairwise comparison of the means for the different factor levels (Field 2009; West et al. 2007). Table 45 shows the results of the post-hoc test, which were calculated by applying the */EMMEANS* sub-command of the SPSS *MIXED* procedure.

Table 44: Parameter Estimates for Fixed Effects (WRONG vs. RIGHT Condition)

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						lower bound	upper bound
Intercept	.653121	.176390	38.699	3.703	.001	.296250	1.009992
[condition=1]	.614281	.271936	40.343	2.259	.029	.064824	1.163738
[condition=2]	0	0	.	.	.	.	.
[phase=1]	.413610	.148377	120.619	2.788	.006	.119850	.707370
[phase=2]	0	0	.	.	.	.	.
[condition=1] * [phase=1]	-.521289	.227117	119.450	-2.295	.023	-.970987	-.071591
[condition=1] * [phase=2]	0	0	.	.	.	.	.
[condition=2] * [phase=1]	0	0	.	.	.	.	.
[condition=2] * [phase=2]	0	0	.	.	.	.	.

For the WRONG condition, the estimate for the average number of snacks increases from 1.209 in the baseline phase to 1.269 in the treatment phase (+5%), which is not significant with a p-value of .741. We therefore reject hypothesis **H1a**, which postulates a decrease of the number of snacks. Under the RIGHT condition, the number of snacks decreases by 0.443 (-40%), which is highly significant with a p-value of 0.005. We therefore accept hypothesis **H1b**. Hypothesis **H1c** is also accepted since the number of snacks under the RANDOM condition decreases by 0.671 (-59%) at a significance level of 0.000.

Table 45: Post-Hoc Test

Condition	Mean (Baseline)	Mean (Treatment)	Mean Difference	Std. Error	df	Sig.	95% Confidence Interval for Difference	
							Lower Bound	Upper Bound
WRONG	1.209	1.269	-.060	.181	168.46	.741	-.418	.298
RIGHT	1.097	.654	.443	.155	168.08	.005	.136	.749
RANDOM	1.141	.470	.671	.165	157.77	.000	.345	.997

To summarize, the LMM analysis has confirmed the expectations we gained from the graphical illustrations in section III.5.3. Whereas under the RIGHT and RANDOM condition we found a significant effect of the persuasive treatment, this effect could not be confirmed for the WRONG group. Although not significant, the number of snacks even increased slightly when non-fitting messages were sent to the subjects. Furthermore, we found that there is no significant difference of the treatment effect between the RIGHT and RANDOM condition. With regard to the initial hypotheses, the LMM analysis has shown that some of the initial expectations could not be confirmed. The hypothesis tests are summarized in Table 46.

Table 46: Summary of Hypothesis Tests

Hypothesis	Description	Test Result
H1a	The number of SNACKS consumed per day is lower in the TREATMENT phase than in the BASELINE phase under the WRONG condition.	rejected
H1b	The number of SNACKS consumed per day is lower in the TREATMENT phase than in the BASELINE phase under the RIGHT condition.	accepted
H1c	The number of SNACKS consumed per day is lower in the TREATMENT phase than in the BASELINE phase under the RANDOM condition.	accepted
H2	Under the WRONG condition, the decrease of the number of SNACKS from the baseline to the treatment phase is lower than under the RANDOM condition.	accepted
H3	Under the RIGHT condition, the decrease of the number of SNACKS from the baseline to the treatment phase is higher than under the RANDOM condition.	rejected
H4	Under the WRONG condition, the decrease of the number of SNACKS from the baseline to the treatment phase is lower than under the RIGHT condition.	accepted

III.6 Discussion

III.6.1 Theoretical Implications

The study has shown that people differ in their level of susceptibility towards different persuasive principles. Applying the most appropriate persuasive strategy can be expected to exert a strong persuasive effect. It was expected that this effect will be significantly weaker when an inappropriate principle is applied. The study has shown that applying an inappropriate persuasive principle may not only weaken the persuasive effect, but can even reverse the effect such that subjects change their behavior in the opposite direction as intended (the results have shown that under the wrong treatment, the number has slightly increased, but not at a significant level). This may indicate that inappropriate treatment causes an aversion towards the desired behavior, which may lead to defiant reactions to act against the extrinsic motivation attempt.

Surprisingly, a random treatment has not shown to be less effective than the most appropriate treatment. Since a random treatment exerts a balanced set of appropriate and inappropriate stimuli, we would have expected that its persuasive effect lies between the right and the wrong treatment. In fact, we even observed that the random treatment performed slightly better than the right treatment, yet not at a significant level. We can only speculate about the underlying reasons for this finding. We assume that not only the appropriateness of a persuasive principle influences its effectiveness, but also the variety of the messages. Applying the same principle several times may lead to an annoyance or boredom effect so that a basically appropriate principle loses its effectiveness if it is applied more often. If this is true, over time the relative advantage of the most appropriate principle decreases, which makes less appropriate principles relatively more effective. Further studies should investigate whether the observed finding that random and right treatments are equivalent can be confirmed, since this will have strong practical implications for the implementation of persuasive principles.

III.6.2 Practical Implications

For the design of persuasive technologies as well as for other disciplines, in which persuasive principles are applied (e.g. marketing or policy), our results have severe practical implications. As applying an inappropriate principle may exert no or even a reverse effect on the subject, single strategy implementations will be ineffective for a part of its users. Two alternatives are useful to consider. First, one could assess which principle is most appropriate for a subject and then apply it. Second one could apply a random selection of principles, which will be equally effective over time.

Selecting the right principle may be an impractical approach for many realistic applications. We have confirmed that it is possible to determine the susceptibility of a subject to the six persuasive principles by applying a questionnaire proposed by Kaptein and Eckles (2010b). Theoretically, we could ask a user of a persuasive application to fill in such a questionnaire before he is exposed to the optimal persuasive treatment, but practically this will be hardly acceptable in many cases. Therefore it would be necessary to develop more acceptable approaches to assess the susceptibility of a user to the different persuasive principles. For PC or mobile applications, it might be possible to achieve an appropriate assessment by designing a game approach. If an application is expected to be used over a longer period of time, it might also be possible to implement a learning algorithm that tests different principles and evaluates their effectiveness before the final principle is selected.

Applying random selection seems to be more applicable in many contexts as it does not require for determining the most appropriate persuasive principle. However, it requires a treatment period that is long enough to apply different principles. If the persuasive treatment is only applied once or twice, many subjects will be exposed solely to inappropriate principles so that the effectiveness may be lower than by applying the most appropriate principle. Instead, if many treatments can be expected, a random selection is preferable as it requires less effort to retrieve a user profile.

III.6.3 Limitations

We identified three limitations of the present study, which leave room for further research approaches. First, the profiling questionnaire has shown reasonable, but not formidable, quality with respect to construct reliability. Particularly, the *Scarcity* construct was at the boundary of acceptability. Further research efforts might lead to a more reliable profiling questionnaire, which will improve the accuracy, with which subjects are assigned to persuasive principles.

Second, the SMS messaging approach taken in this study to apply persuasive principles is limited, as only four out of the six persuasive principles identified by Cialdini (2001) could be implemented. Different experimental designs should be developed to include all six principles. This might be particularly valuable to further investigate our observation that right and random treatments are equivalent.

Third, our sample size was limited so that our results are valid for our sample, but may lack of generalizability to the population. Furthermore, we propose larger samples to investigate the insignificant observation that an inappropriate treatment may have a reverse effect, and that random and right treatments are equivalent.

IV Behavior-based Business Models

IV.1 Case Description

The goal of the study presented in this chapter is to answer the third research question raised in this dissertation:

Q3: *Which factors influence the acceptance of behavior-based automobile insurance and their effect on the willingness to change driving behavior?*

The previous two studies investigated persuasive approaches that intend to influence human behavior without monetary incentives. In this third study, we analyze how people perceive business models that provide financial rewards for adopting a desired behavior. Furthermore, we investigate whether consumers who decide to accept persuasive business models can be expected to change their behavior.

In many domains of daily life, a third party has an interest that people behave in a certain way. For example, health insurance companies have an interest that their customers engage in physical workout, do not smoke or reduce instances of being overweight. Similarly, automobile insurances and automobile leasing companies are interested in motivating their customers to drive cautiously and treat their vehicles carefully. Sensor technologies and ubiquitous connectivity allow for an accurate monitoring of habits and behaviors in many domains, which reveals the possibility to reward desired or punish undesired behavior. We define business models that reward or punish certain behavioral patterns as *behavior-based business models*.

As a concrete example for behavior-based business models, we investigate how consumers perceive a behavior-based automobile insurance model. Under such an insurance contract, the premium depends on driving behavior. A technical device captures a number of driving parameters, which are evaluated to infer the accident risk a driver may have. If a high accident risk is estimated, the policy holder has to pay a larger premium than if a low risk is expected.

The simplest form of behavior-based insurance is the "pay-as-you-drive" model, under which the insurance premium depends on the driven distance (Bordoff and Noel 2008). Mileage can either be read manually or transmitted online via a mobile phone network. An alternative is a surcharge on fuel prices, in literature often referred to as "pay-at-the-pump" (Litman 2008). Online implementations of pay-as-you-drive insurance are offered by several insurance companies, among them Aioi (Japan), Amaguiz (France), GMAC (USA), Progressive (USA), Real Insurance (Australia) and Uniqa

(Austria) (Ippisch 2010). These offerings require installing a device that transmits travelled distances to the insurance company via the mobile phone network. Travelled distances are either obtained via GPS or by connecting the device to the on-board diagnostic system.

More advanced offerings take into account not only mileage but also driving conditions like daytime or road profile. The rationale is that driving in the night or on crowded roads increases accident risk. Acorn (UK), Aviva (France), Axa (Italy), and Solly Azar (France) are examples for insurance companies that offer such policies. Coverbox (UK) acts as an intermediary as they provide the technology to the customer, who can then choose among several collaborating insurers who offer distance-based insurance contracts (e.g. Allianz, Sabre, Groupama).

Most sophisticated behavior-based insurance models additionally evaluate driving behavior itself by assessing parameters like speed, acceleration, deceleration or curve speed (Toledo et al. 2008). Based on these measurements, a risk index is calculated that influences the insurance premium. Insurance companies Liberty Mutual (USA), Progressive (USA), and WGV (Germany) offer such behavior-based insurance models in the narrow sense, which are the object of investigation in the present study. In the following section we summarize research work that is relevant in the context of behavior-based automobile insurance.

IV.2 Related Work

Three different research streams can be identified for contributing to the analysis of behavior-based automobile insurance models from an economic perspective. First, there is a large body of work on information asymmetries in insurance markets. A second stream investigates the socio-economic impact of behavior-based vehicle insurance, and a third research direction investigates barriers to a wider adoption of these insurance models. In the following, major findings from these three research streams are summarized.

IV.2.1 Information Asymmetries in Insurance Markets

Before an insurance contract is closed, information asymmetries arise from the fact that the insurance applicant has more complete information about his risk profile than the insurance company. Akerlof (1970) argues that these information asymmetries lead to market inefficiencies and an unbalanced risk portfolio due to *adverse selection*, which is "potentially present in all lines of insurance".

“Adverse selection can be defined as the process by which prospective policyholders may gain a financial advantage through insurance purchase decisions based on risk characteristics known to them, but unknown and not revealed to the insurer” (Subramanian et al. 1999). Under the assumption that individuals can self-assess their accident risks, adverse selection can occur in automobile insurance markets when high-risk drivers chose complete coverage, whereas low-risk drivers decide for a policy with partial coverage. Rothschild and Stiglitz (1976) have shown that under this assumption, adverse selection will impede optimal market equilibrium as high-risk drivers will exert negative externalities on the low-risk group. They conclude that "low-risk individuals are worse off than they would be in the absence of the high-risk individuals", whereas high-risk individuals do not benefit from the presence of low-risk individuals. The presence of adverse selection in automobile insurance markets was empirically confirmed by Dahlby (1983) and D’Arcy and Doherty (1990).

After an insurance contract is closed, the lack of information about actual behavior may lead to *moral hazard*, i.e. "the tendency of insurance protection to alter an individual's motive to prevent loss" (Shavell 1979). Two general strategies are proposed to mitigate moral hazard issues. Either part of the financial risk stays with the policy holder as an incentive to prevent loss, or insurance conditions are linked to observed behavior so that loss prevention leads to lower premiums or higher coverage (Arrow 1963; Pauly 1968). Filipova-Neumann and Welzel (2010) investigated the effects of black-boxes that monitor driving-behavior, which is evaluated in case of an accident. Applying the insurance model from Rothschild and Stiglitz (1976), they have shown that under perfect competition, the proposed black-box approach will mitigate the adverse effects of information asymmetries and lead to a pareto-improvement of social welfare. They thus support the general conclusion from Rothschild and Stiglitz (1976) for the concrete case of a vehicle black-box, that "if individuals were willing or able to reveal their information, everybody could be made better off". It may further mitigate the problem of so-called *statistical discrimination*, which arises from today's practice to estimate accident risks primarily on the basis of demographic information (Dahlby 1983). Although this approach may lead to acceptable insurance conditions for a majority of policy holders, a substantial share of insurance customers can be expected to be better off if their accident risk is assessed on the basis of actual behavior instead of demographic patterns.

IV.2.2 Socio-economic Impact

A number of studies investigate the societal impact of behavior-based automobile insurance with regard to fuel consumption, accident rates, and emissions. Vickrey (1968) was one of the first to analyze the societal impact of behavior-based automobile insurance. He proposed a mileage-based premium scheme to overcome the issues that lump sum schemes raise, since they neither account for external costs of accidents nor provide incentive for a desirable change in behavior. In the absence of affordable telematics technologies at the time his analysis was conducted, he proposed to tie insurance premiums on gasoline sales so that higher mileage leads to higher insurance premiums. He acknowledged that mileage is not the only factor to influence accident risk. To link insurance premiums to driving behavior, he therefore proposed to charge a surplus on tire prices, which would result in higher insurance premiums for tire-abrasive driving style. Consequently, drivers would be motivated not only to drive less but also to adopt a more careful driving behavior.

Parry (2005) analyzes the potential of mileage-based insurance schemes ("pay-as-you-drive") to lower fuel demand and increase social welfare. He concludes that although increased gasoline taxes would be an appropriate means to reduce fuel consumption and emissions, a mileage-based insurance scheme is superior with regard to social welfare. He argues that increased gasoline prices can be compensated by switching to more fuel-efficient vehicles. In contrast to that, distance-based insurance premiums can only be influenced by reducing mileage. As a consequence, they do not only lead to reduced fuel consumption and emissions, but have an additional positive influence on the number of accidents and congestion.

Varying estimations have been proposed about the fuel saving potential of mileage-based insurance schemes. Parry (2005) concludes that such an insurance scheme would reduce fuel demand by 9.1%, whereas Bordoff and Noel (2008) estimate a saving potential of 2%. Edlin (2003) estimates that pay-as-you-drive insurance would reduce annual mileage in the USA by 9.5%.

Investigating the effect of behavior-based insurance models on accident rates, Litman (2008) assumes that a 1% reduction in total vehicle mileage will reduce the number of crashes by 1.2%. If GPS-based pricing is applied to give an incentive to avoid high-risk driving conditions, a 1% mileage reduction may even lead to a crash reduction 1.4-1.6%. Toledo et al. (2008) investigate how individual feedback on driving behavior influences driving style. They show that a risk index they developed to measure the risk-proneness of an individual's driving behavior based on in-vehicle sensor data had significantly decreased after providing feedback. Zantema et al. (2008) conducted a simulation study to compare seven different mileage-based insurance schemes with regard to their impact on congestion avoidance and accident reduction. They found that different insurance models have to be implemented to optimize either congestion or accident rates. A crash reduction of 5% is achievable if insurance premiums are based on mileage, driving time and road type.

Taking into account externalities from decreasing accident rates and congestion, Edlin (2003) developed an analysis model to quantify the monetary effect of completely switching to mileage-based insurance models. He estimates external accident costs to be 0.075 USD per mile, whereas insurance and gasoline costs lie between 0.04 and 0.06 USD. The difference is imposed as social costs on society. He concludes that a monetary benefit between 76 and 108 USD per vehicle and year can be achieved through mileage-based insurance (neglecting monitoring costs) through reduced accident rates and related externalities, and through reduced congestion costs due to reduced traffic. Bordoff and Noel (2008) even estimate a monetary benefit of 257 USD per vehicle and year. Insurance-related savings in their comprehensive calculation base consist of external (93 USD) and individual (34 USD) auto insurance cost savings. Total savings also include a monetary value for reduced congestion (58 USD), saved carbon emissions (11 USD), reduced oil consumption (25 USD) and other positions. Therefore, only 127 US of the total benefit may be returned to the customer by the insurance company. The majority of the savings, however, is achieved through reduced externalities, which may increase social welfare but falls out of the scope of an insurance business model. Edlin (2003) supports this view by noting that the majority of benefits will not remain at the insurance companies that introduce mileage-based models as they are primarily gained through externalities (i.e. the reduction of insurance, accident and congestion costs of others), which may hinder insurance companies to promote mileage-based insurance models.

IV.2.3 Barriers to Adoption

Despite the positive benefits which behavior-based automobile insurance is expected to provide, a broad adoption has not yet been taken place. Monitoring costs have long dominated the discussion about barriers to a successful introduction of behavior-based insurance models (Bordoff and Noel 2008; Edlin 2003; Rea 1992; Vickrey 1968). However, the increasing proliferation of telematics solutions in vehicles can be expected to propel a further decline in monitoring costs once a majority of vehicles are equipped with online telematics platforms that allow for the implementation of third-party applications (Ippisch 2010). Such platforms would abandon the need for dedicated devices to monitor driving behavior. Instead, a piece of software could be installed on the telematics platform to monitor driving behavior and send this information to the insurance company.

Bordoff and Noel (2008) identify regulatory burdens and patent disputes as two other barriers of pay-as-you-drive insurance. Referring to the situation in the United States, the authors point out that in several U.S. states mileage-based insurance premiums would not be legal under current law. Guensler et al. (2003) analyzed the legal situation in 43 U.S. states. 16 states would not allow mileage-based insurance premiums. The remaining 27 states impose a number of regulatory burdens to ensure equity and transparency of the pricing structure.

Furthermore, Bordoff and Noel (2008) point out that key patents required to implement behavior-based insurance models are held by the U.S. insurance company Progressive. Consequently, insurance companies willing to adopt a behavior-based insurance model may be concerned about patent infringements.

Comparably weak attention has been paid to barriers that potentially arise from customers' attitudes towards certain aspects of behavior-based insurance models. Several authors mention privacy concerns as a potential impediment for a broad adoption (Bordoff and Noel 2008; Lindberg et al. 2005; Litman 2008), but their concrete effect as a barrier for consumer acceptance has hardly been analyzed. This research gap is intended to be filled by the present study.

IV.3 Research Design

Similarly to our analysis in section II, we developed a research model on the basis of the *Unified Theory of Acceptance and Use of Technology* (UTAUT). However, since behavior-based insurance models represent not only a technical innovation, but extend traditional insurance contracts by a novel pricing scheme, technical aspects cannot be expected to account for all relevant factors that influence adoption decisions. At the same time, we identified a lack of appropriate research models in the existing body of work, which would cover both the technical and non-technical dimensions of behavior-based insurance. Consequently, we extended the UTAUT model by adding a number of constructs that turned out to influence adoption decisions in prior research work. The constructs were arranged and re-interpreted such that they capture aspects that relate to both the underlying technology as well as the proposed pricing scheme. As a result, we propose a new research model that extends prior technology acceptance models by a business model dimension. We explicitly underpin the exploratory nature of our analysis since our goal is not to confirm an existing research model, but to develop and test a novel structural model that accounts for the idiosyncrasies of behavior-based business models in order to uncover relevant factors that influence consumer perception.

Existing technology acceptance research usually regards the behavioral intention to adopt a proposed technology as the only dependent variable concerning *Behavioral Intention*, which serves as a predictor of actual use (Venkatesh et al. 2003). In the context of behavior-based business models, we are not only interested in the intention to adopt the proposed offering, but also in the intention to change current behavior. We assume that the intention to adopt a behavior-based business model will influence the intention to change behavior. Therefore we added a new construct *Behavioral Intention to Change* (BIC). In our concrete context, this construct describes the intention of a person to change driving behavior under a behavior-based automobile insurance contract. Furthermore, we assume that the relationship between the *Behavioral Intention to Adopt* (BI) and the *Behavioral Intention to Change* is moderated by *Cost Sensitivity*, *Driving Style* and *Gender* (Figure 14).

In accordance with existing literature, we included *Attitude* towards the proposed insurance model as a predictor of *Behavioral Intention to Adopt* (Ajzen 1991; Nysveen et al. 2005; Schepers and Wetzels 2007). From the UTAUT model, we employed *Performance Expectancy, Effort Expectancy* and *Social Influence* as antecedents of both *Attitude* and *Behavioral Intention* (Venkatesh et al. 2003). *Perceived Enjoyment* was added to the model to capture the hedonic dimension of the proposed insurance model and its influence on *Attitude* and *Behavioral Intention to Adopt* (Davis et al. 1992). Privacy concerns have been shown to influence user acceptance of online services (Chellapa and Sin 2005; Dinev and Hart 2006; Sheng et al. 2008; Shin 2009). We therefore included *Perceived Privacy* as a predictor to *Attitude* and *Behavioral Intention to Adopt*. Finally, we added *Trust* in the reliability and honesty of the insurance provider as an antecedent to the *Behavioral Intention to Adopt*, which has been confirmed by a large number of technology acceptance studies (Gefen and Straub 2004; McKnight et al. 2002; Pavlou and Gefen 2004; Shin 2009).

Figure 14 shows the research model developed for the present study. In the following, we will define the different constructs, explain how we adapted them to the given context, and introduce the hypothesized relationships among them.

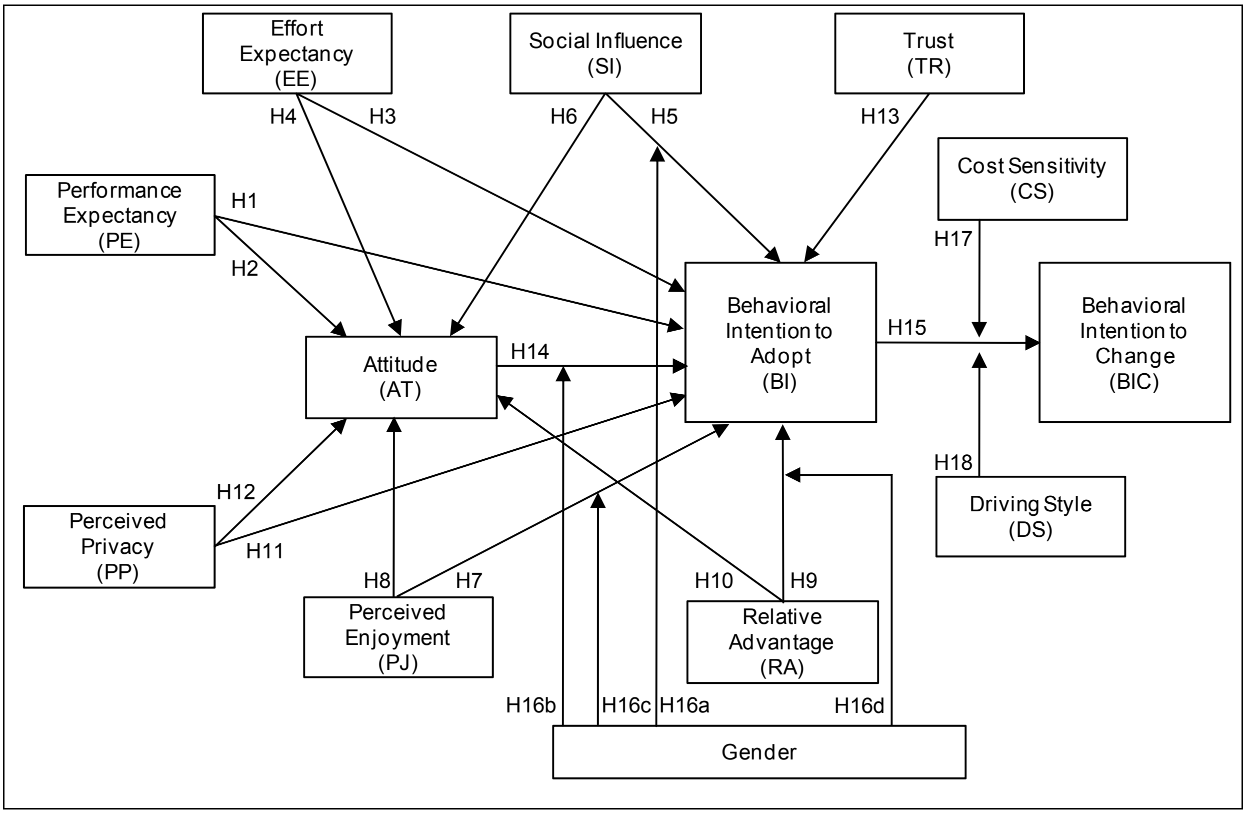


Figure 14: Research Model

Performance Expectancy (PE)

Performance Expectancy (PE) has repeatedly been confirmed to be the strongest predictor for the behavioral intention to use information systems (Schepers and Wetzels 2007; Venkatesh et al. 2003). In the context of IS research, it has been defined as "the degree to which an individual believes that using the system will help him or her to attain gains in job performance" (Venkatesh et al. 2003). In our context, technology does not directly produce benefits for the consumer but is limited to enabling the implementation of the proposed insurance model. We therefore adapted the PE construct such that it measures the expectancy that the underlying technology will monitor and evaluate an individual's driving behavior correctly. In accordance with previous work, we assume that PE has a positive influence on the *Behavioral Intention to Adopt* as well as on the general *Attitude* towards the proposed insurance model (Nysveen et al. 2005; Venkatesh et al. 2003). We therefore hypothesize:

> **H1:** *Performance Expectancy* has a positive effect on *Behavioral Intention to adopt*.
>
> **H2:** *Performance Expectancy* has a positive effect on *Attitude*.

Effort Expectancy (EE)

Venkatesh et al. (2003) define *Effort Expectancy* (EE) as "the degree of ease associated with the use of the system". EE aggregates the three related constructs *Perceived Ease of Use*, *Complexity,* and *Ease of Use* from earlier technology acceptance models, which describe the degree to which using a system is perceived as being "free of effort" or "difficult to understand and use" (Davis 1989; Moore and Benbasat 1991; Thompson et al. 1991). In our context, potential consumers do not directly interact with the proposed technology or have to change their behavior. However they might perceive it as effortful to understand the proposed business model, control their driving behavior and get used to varying insurance premiums. We therefore adapted the EE construct such that it represents the perceived effort to deal with the novelty and complexity of the proposed insurance model. In accordance with previous work, we assume that *Effort Expectancy* has a positive effect on the *Behavioral Intention to Adopt* (Venkatesh et al. 2003) and on *Attitude* (Nysveen et al. 2005; Schepers and Wetzels 2007; Taylor and Todd 1995). Hence, we hypothesize:

> **H3:** *Effort Expectancy* has a positive effect on *Behavioral Intention to adopt*.
>
> **H4:** *Effort Expectancy* has a positive effect on Attitude.

Social Influence (SI)

Social Influence (SI) has been defined as "the degree to which an individual perceives that important others believe he or she should use the new system" (Venkatesh et al. 2003). SI is based on the *Social Norm* construct of the *Theory of Planned Behavior*, which is defined as "the person's perception that most people who are important to him think he should or should not perform the behavior in question" (Fishbein and Ajzen 1975). Following these two definitions, we relate the SI construct to the question of whether important others would appreciate that a person accepts the proposed insurance model. In accordance with previous findings in information systems research, we assume that SI has a positive effect on BI (Hong et al. 2008; Nysveen et al. 2005; Schepers and Wetzels 2007; Venkatesh et al. 2003). We further postulate that, in accordance with the *Theory of Planned Behavior,* but unlike many information systems studies, SI has also a positive influence on *Attitude* (Ajzen 1991). The reason for this assumption is that it might be rather difficult for people to evaluate cognitively the proposed business model in its whole complexity. In such a situation, people tend to refer to heuristic evaluation approaches such as relying on social norms in order to overcome informational uncertainties and to avoid excessive cognitive effort. Therefore we assume that *Social Influence* does not only influence BI but also the general attitude towards the proposed insurance model. Hence, we hypothesize:

H5: *Social Influence* has a positive effect on *Behavioral Intention to Adopt*.

H6: *Social Influence* has a positive effect on *Attitude*.

Perceived Enjoyment (PJ)

Davis et al. (1992) introduced *Perceived Enjoyment* (PJ) as an antecedent to *Behavioral Intention* into their technology acceptance model TAM. They define *Perceived Enjoyment* as "the extent to which the activity of using the computer is perceived to be enjoyable in its own right, apart from any performance consequences that may be anticipated". Several studies have shown that PJ has a weak but significant effect on information system acceptance in a professional setting, particularly if only moderate productivity gains are expected (Davis et al. 1992; Igbaria et al. 1996, 1994). In contrast to the results for utilitarian systems, the purpose of which is to provide instrumental value, PJ has been found to be a dominant predictor for behavioral intention with respect to hedonic systems like computer games or web applications (Atkinson and Kydd 1997; Venkatesh 1999; Moon and Kim 2001; Heijden 2004). These results are consistent with findings from the consumer behavior literature, which point out that consumer decisions may be guided either by utilitarian or by hedonic considerations (Holt 1995; Venkatraman and MacInnis 1985). With regard to the insurance model under investigation in this study, we expect only a minor hedonic value of the insurance model itself. However, it may impact the perceived enjoyment of driving a vehicle. Persons who prefer a driving style that is disadvantageous from an insurer's perspective may feel a loss of driving pleasure. Furthermore, the behavior-based insurance may cause anxiety about uncontrollable cost consequences. We therefore deviate from the original definition of the PJ construct such that we do not measure the perceived enjoyment experienced by the insurance model itself, but by measuring its expected consequences for driving pleasure. We assume that PJ positively influences the *Behavioral Intention to Adopt* (Davis et al. 1992; Heijden 2004) and *Attitude* (Hong et al. 2008; Moon and Kim 2001; Nysveen et al. 2005). Hence, we hypothesize:

H7: *Perceived Enjoyment* has a positive effect on *Behavioral Intention to Adopt*.

H8: *Perceived Enjoyment* has a positive effect on *Attitude*.

Relative Advantage (RA)

Relative Advantage (RA) has been determined by Rogers (1983) as one out of five attributes that influence the adoption of innovations. He defined *Relative Advantage* as "the degree to which an innovation is perceived as being better than its precursor". Among others, Moore and Benbasat (1991) and Premkumar and Potter (1995) have confirmed RA as a predictor to the intention to use an information system. Venkatesh et al. (2003) have subsumed *Relative Advantage* under the *Performance Expectancy* construct in their UTAUT model. In our model, we intend to separate functional performance aspects from the benefits expected by individuals, which are primarily monetary rewards for careful driving. However, since individuals might also expect non-monetary benefits like education to careful driving, we do not limit RA to monetary aspects. We therefore re-interpret the RA construct such that it describes the degree to which the proposed insurance model is expected to provide a personal benefit. Since RA represents a specialization of the *Performance Expectancy* construct, we assume the same effects as for PE. We hypothesize:

> **H9:** *Relative Advantage* has a positive effect on *Behavioral Intention to Adopt*.
>
> **H10:** *Relative Advantage* has a positive effect on *Attitude*.

Perceived Privacy (PP)

Privacy is the "the ability to control the terms by which personal information is acquired and used" (Westin 1967). Privacy concerns arise from the possibility that companies collect data and use them inappropriately (Jarvenpaa and Toad 1996; Roca et al. 2009). *Perceived Privacy* (PP) depends on technical and non-technical considerations (Mukherjee and Nath 2007; Roca et al. 2009). The technical infrastructure has to ensure that user data is only accessible to authorized persons and organizations. In literature, this aspect is often referred to as "security" (Mukherjee and Nath 2007; Roca et al. 2009) or "perceived security" (Shin 2009). Furthermore, organizational measures are necessary to ensure that retrieved user data are treated confidentially (Mukherjee and Nath 2007). On the basis of the explanations from Jarvenpaa and Toad (1996) and Mukherjee and Nath (2007), we define *Perceived Privacy* as the degree to which the proposed business model is perceived to ensure that personal data is not used inappropriately. The proposed insurance model will collect a large amount of critical personal data such as information about an individual's driving style, travel routes, visited places or adherence to speed limits. We therefore assume that privacy concerns influence both *Attitude* and *Behavioral Intention to Adopt*. In accordance with previous work, we hypothesize (Chellapa and Sin 2005; Dinev and Hart 2006; Sheng et al. 2008; Shin 2009):

H11: *Perceived Privacy* has a positive effect on *Behavioral Intention to Adopt*.

H12: *Perceived Privacy* has a positive effect on *Attitude*.

Trust (TR)

Depending on research discipline and research goals, various different definitions for trust have been suggested, which lead to Shapiro's conclusion that trust definitions are a "confusing pot-pourri". With respect to IS research, trust is often associated or mixed with privacy and security considerations (Mukherjee and Nath 2007; Roca et al. 2009; Shapiro 1987), which are captured by the *Perceived Privacy* construct in our research model. In the present study, we regard *Trust* as "the confident truster perception that the trustee […] has attributes that are beneficial to the truster" (McKnight et al. 2002). Three so-called "trusting beliefs" primarily contribute to this kind of trust (McKnight et al. 2002): "competence (ability of the trustee to do what the truster needs), benevolence (trustee caring and motivation to act in the truster’s interests), and integrity (trustee honesty and promise keeping)". In our concrete context, the *Trust* construct measures the degree to which a person believes that an insurance company has the competence to correctly evaluate driving data, and that it is credible and trustworthy enough to assume that the proposed insurance model would be a fair offering. As we relate *Trust* to the institution of an insurance company and not to the proposed insurance model in general, we hypothesize that *Trust* only influences BI but not AT (Gefen and Straub 2004; Pavlou and Gefen 2004; Shin 2009):

H13: *Trust* has a positive effect on *Behavioral Intention to Adopt*.

Attitude (AT)

Attitude has been introduced by Fishbein and Ajzen (1975) as a predictor to behavioral intention in the Theory of Reasoned Action (TRA), and later in the Theory of Planned behavior (TPB) (Ajzen 1991). It has been defined as "the degree to which a person has favorable or unfavorable evaluation or appraisal of the behavior in question" (Ajzen 1991). Although the technology acceptance models TAM, TAM2 and UTAUT are based on the ideas of TRA and TPB, *Attitude* has been excluded from these models to achieve a more parsimonious model (Venkatesh et al. 2003). Nevertheless, many empirical studies have shown that *Attitude* has a significant mediating role to predict the intention to adopt a certain technology. Some authors suggest to include *Attitude* as a full mediator, which means that antecedents of attitude (e.g. effort expectancy or performance expectancy) have no direct relationship to behavioral intention but are fully mediated by *Attitude* (Dabholkar and Bagozzi 2002; Hong et al. 2008; Shin 2009). Other authors model *Attitude* as a partial mediator, which means that antecedents of *Attitude* have a direct effect on behavioral intention and are at the same time partially mediated via *Attitude* (Moon and Kim 2001; Nysveen et al. 2005; Schepers and Wetzels 2007; Taylor and Todd 1995). Consumer behavior research confirms the important role of attitude as a mediator of antecedents like product price, quality, involvement and emotion (Hansen 2005). Against the background of strong empirical support for the explanatory power of attitude, we include *Attitude* as a mediator for SI, EE, PE, PP, PJ, and RA into our research model. We apply *Attitude* as a measure for the degree to which a person perceives the concept of behavior-based automobile insurance as a favorable and desirable idea. We hypothesize:

H14: *Attitude* has a positive effect on *Behavioral Intention to Adopt*.

Behavioral Intention to Adopt (BI)

Consistent with the Theory of Reasoned Action, Theory of Planned Behavior, and various technology acceptance models, we assume that behavioral intention is a robust predictor of actual behavior (Ajzen 1991; Fishbein and Ajzen 1975; Venkatesh et al. 2003). We therefore measure the *Behavioral Intention to Adopt* (BI) the proposed insurance model in order to judge the degree to which our subjects are willing to prefer a behavior-based automobile insurance over a conventional one. Originally named *Behavioral Intention* (Venkatesh et al. 2003), we renamed this construct into *Behavioral Intention to Adopt* in order to distinguish it from *Behavioral Intention to Change*. We extend existing technology acceptance models by exploring the assumption that *Behavioral Intention to Adopt* has a positive influence on the willingness of a person to change his driving behavior. The purpose of the underlying hypothesis is to test whether behavior-based business models may represent an appropriate means to influence the behavior of people towards a desirable direction. Hence we hypothesize:

H15: *Behavioral Intention to Adopt* has a positive effect on *Behavioral Intention to Change*.

Behavioral Intention to Change (BIC)

Behavioral Intention to Change (BIC) has been added to our research model as a new construct in order to explore whether people who are willing to adopt the proposed insurance model are also willing to change their driving style. The goal of this exploratory evaluation is to predict whether behavior-based automobile insurance models may contribute to motivate drivers to sustain a safer and more careful driving style. We define *Behavioral Intention to Change* as the degree to which a person is willing to adopt a more careful driving behavior in the presence of a behavior-based automobile insurance contract.

Moderating Effects

Studies investigating gender differences in driving behavior have shown that male drivers drive generally more than female drivers, speed more often, are more inclined to drive under the influence of alcohol, and exhibit more dangerous thought patterns with respect to road traffic (Harré et al. 1996). Men are additionally more likely to exert highly aggressive and violent behavior, particularly under conditions of stress (Hennessy and Wiesenthal 2001). Female drivers have a higher risk perception with regards to driving, whereas men have a generally more positive affect towards driving and a higher affinity to adopt a risky driving style (Rhodes and Pivik 2011). Taking into account the broad empirical evidence for gender differences in driving behavior and risk perception, we assume that gender significantly moderates the relationships between *Behavioral Intention to Adopt* and its antecedents.

Psychological studies investigating the disposition to comply with social norms have shown that women are more inclined to comply with orders (Minton et al. 1971; Minton and Schneider 1980; Stockard et al. 1988), to conform with a majority opinion (Eagly 1978; Maccoby and Jacklin 1974) and to show more submissive character traits (Crawford et al. 1995), making them more susceptible to social cues and information received from others (Garai and Scheinfeld 1968; Roberts 1991). In accordance with these findings, technology acceptance studies have shown that women are more strongly influenced by *Social Influence*, whereas men base their technology usage decisions on *Attitude* and *Performance Expectancy* (Venkatesh et al. 2003, 2000; Venkatesh and Morris 2000). We therefore assume that the effect of *Social Influence* is larger for women than for men, whereas men are more influenced by *Attitude* than women. In the original UTAUT model, aspects captured by our *Relative Advantage* construct were partially subsumed under *Performance Expectancy*, which would induce the assumption that RA has a stronger influence for men than for women. However since women usually drive more carefully and less aggressive, they can be expected to experience a larger relative advantage than men. Therefore, we assume that, in contrast to other technology acceptance studies, *Relative Advantage* will have a stronger influence for women than for men in our specific context. Following this line of argumentation, we also expect that *Perceived Enjoyment* will exert a stronger influence for women than for men because the relative loss in driving pleasure is expected to be lower for women than for men.

H16a: For women the effect of *Social Influence* on *Behavioral Intention to Adopt* is higher than for men.

H16b: For men the effect of *Attitude* on *Behavioral Intention to Adopt* is higher than for men.

H16c: For women the effect of *Perceived Enjoyment* on *Behavioral Intention to Adopt* is higher than for men.

H16d: For women the effect of *Relative Advantage* on *Behavioral Intention to Adopt* is higher than for men.

With the *Behavioral Intention to Change* construct, we added a theoretical extension to the empirically confirmed basis of our research model. The BI-BIC relationship reflects our assumption that the intention to change one's own driving style increases with the intention to accept behavior-based automobile insurance. We further assume that this relationship is moderated by the two variables *Cost Sensitivity* and *Driving Style*. With *Cost Sensitivity,* we describe the degree to which a person perceives it to be important to keep the total costs of owning a vehicle at a low level. We assume that cost-sensitive people are more willing to adapt their driving style in exchange for a monetary reward than people who are less concerned about the costs for their automobile. *Driving Style* measures the degree to which a person drives carefully and cautiously. Cautious drivers have a lower incentive to adapt their driving style under a behavior-based insurance contract than sportive or aggressive drivers. We therefore expect that cautious drivers are less willing to change their driving behavior than reckless and aggressive drivers. For both variables, we assume that they have a moderating effect on the BI-BIC relationship as well as a direct effect on BIC. Hence we hypothesize:

H17a: *Price Sensitivity* has a positive effect on *Behavioral Intention to Change*.

H17b: The effect of *Behavioral Intention to Adopti*on in *Behavioral Intention to Change* increases with higher *Price Sensitivity*.

H18a: *Driving Style* has a negative effect on *Behavioral Intention to Change*.

H18b: The effect of *Behavioral Intention to Adopti*on on *Behavioral Intention to Change* decreases with higher levels of *Driving Style*.

IV.4 Data Collection

IV.4.1 Instrument Development

The measurement instruments used in this study were based on the existing body of work in the field of technology acceptance research. But, because these instruments were developed for assessing the perception of technical innovations, adaptations had to be made for the evaluation of a technology-based business model. The questionnaire used in this study was originally formulated in German. Table 47 shows the English translation of the German questionnaire items. All items were measured on 7-point Likert scales or 7-point semantic scales. The Likert scales ranged from "strongly disagree" to "strongly agree". For the semantic scales, the scale endpoints are shown in Table 47.

Items for the *Attitude* construct were based on Fishbein and Ajzen (1975) and Taylor and Todd (1995). AT2 and AT3 were newly developed to capture the respondents' attitude with respect to safety and environmental concerns. Items for *Behavioral Intention to Adopt* were based on (Venkatesh et al. 2003). The concrete formulation had to be changed as the insurance offering is not yet available. Measurement scales for *Behavioral Intention to Change* were newly developed. Measurements for *Trust* were adapted from McKnight et al. (2002) and Gefen et al. (2003). *Effort Expectancy* scales were developed in accordance with Venkatesh et al. (2003) and Moore and Benbasat (1991). In contrast to the original scales, we did not relate *Effort Expectancy* to the use of the technology but to the understanding and handling of the proposed insurance contract. The *Performance Expectancy* items, which were developed on the basis of Venkatesh et al. (2003), were related to the respondents' expectation that the proposed technical approach will be capable of correctly evaluating driving behavior. Existing scales for *Perceived Enjoyment* - as described for example by Davis et al. (1992) or Heijden (2004) - measure the hedonic value of technical innovations. The hedonic value of behavior-based automobile insurance is captured by PJ4. The other PJ items were newly developed to measure the degree to which such an insurance model is expected to influence driving pleasure. *Perceived Privacy* measurements were developed on the basis of Dinev and Hart (2006), Mukherjee and Nath (2007) and Sheng et al. (2008). Scales for *Relative Advantage* described in previous work typically evaluate expected efficiency gains through technical artifacts or their capability to accomplish the intended task (Moore and Benbasat 1991; Premkumar and Potter 1995). In the given context however, *Relative Advantage* is primarily gained from monetary rewards for adopting a cautious driving style, by increasing the perceived fairness of insurance premiums, and by getting an evaluation of one's own driving style. Items were developed to measure the perceived superiority of the proposed insurance model with regard to these aspects. *Social Influence* was measured based on the items from Venkatesh et al. (2003) and Hong et al. (2008). *Cost Sensitivity* items were developed to measure the degree to which a person tries to hold down costs for driving a vehicle (CS1) and perceives driving as expensive (CS2, CS3). The *Driving Style* measure for risk-free driving behavior has been evaluated on the basis of three newly developed items that ask for keeping an accurate distance to other vehicles (DS1), for adopting a generally careful and cautious driving style (CS2), and for compliance with speed limits. To summarize, the majority of items were based on existing measurement scales from technology adoption research. Due to the fact, however, that the present study evaluates a novel business model based on a technical innovation, adaptations had to be made to meet the requirements of the given context.

Table 47: Measurement Instruments

Construct	Item	Question
Attitude (AT)	AT1	I perceive the described behavior-based automobile insurance model as a… (bad / good idea).
	AT2	I think that behavior-based automobile insurances would make traffic safer.
	AT3	I think that behavior-based automobile insurances would motivate to adopt an ecological driving style.
	AT4	I perceive the described behavior-based automobile insurance model as… (undesirable / desirable).
	AT5	Towards the described behavior-based automobile insurance model, I have a… (bad / good feeling).
Behavioral Intention (BI)	BI1	Once offered, I would opt for behavior-based automobile insurance in future.
	BI2	I would allow my insurance company to evaluate my driving behavior if I could thereby achieve a lower insurance rate.
	BI3	I wish that the introduced offering will be available soon.
	BI4	I would accept insurance rate that depend on my driving behavior.
Behavioral Intention to Change (BIC)	BIC1	I would change my driving style if I would thereby pay a lower insurance rate.
	BIC2	Behavior-based automobile insurance would motivate me to drive more carefully and cautiosly.
	BIC3	Behavior-based automobile insurance would influence my driving style positively.
Trust (TR)	TR1	I would trust my insurance company to evaluate my driving behavior correctly.
	TR2	I would trust my insurance company that this kind of insurance contract is a fair offering.
	TR3	I consider my insurance company as credible.
Effort Expectancy (EE)	EE1	It is easy for me to understand which consequences behavior-based automobile insurance would have for me.
	EE2	It would be easy for me to deal with the terms of contract of behavior-based automobile insurance.
	EE3*	Behavior-based automobile insurance would be confusing for me.
	EE4*	I would require a long time to get used to behavior-based automobile insurance.

Performance Expectancy (PE)	PE1	I have confidence in the technology that it is possible to evaluate driving behavior correctly.
	PE2	I believe that it is possible to evaluate the risk of getting involved into an accident by observing a driver's behavior.
	PE3	I would trust a technology for evaluating driving behavior to work reliably.
Perceived Enjoyment (PJ)	PJ1*	The proposed behavior-based automobile insurance contract would lower my driving pleasure.
	PJ2*	With a behavior-based automobile insurance I could not enjoy driving any more.
	PJ3*	Driving pleasure would be afflicted under a behavior-based insurance contract.
	PJ4	It would be pleasant for me if I could influence my insurance rate through my driving style.
Perceived Privacy (PP)	PP1	I would have confidence in my insurance company to treat my driving data confidentially.
	PP2	I trust technology that my driving data won't fall into wrong hands.
	PP3*	Under such a contract I would have the feeling to be observed.
Relative Advantage (RA)	RA1	Behavior-based automobile insurance would be beneficial for me.
	RA2	I think that behavior-based automobile insurance would safe me some money in comparison to a normal contract.
	RA3	With a behavior-based automobile insurance, I would perceive my insurance rates to be fairer.
	RA4	I would perceive it as useful to regularly get an accurate analysis of my driving behavior.
Social Influence (SI)	SI1	People who are important to me would appreciate me to close a behavior-based automobile insurance contract.
	SI2	People who are important to me would advise me to close a behavior-based automobile insurance contract.
	SI3	People who are important to me would regard behavior-based automobile insurance as a good idea.
Cost Sensitivity (CS)	CS1	It is important to me to drive particularly inexpensively.
	CS2	I perceive the total costs for a vehicle as very high.
	CS3	A vehicle is a large cost factor for me.
Driving Style (DS)	DS1	I usually hold an accurate distance to other vehicles.
	DS2	I usually drive carefully and cautiously.
	DS3	I usually adhere to speed limits.
Items marked with (*) were measured on a reverse scale. If no scale is indicated, a 7-point Likert scale was applied, which ranged from "strongly disagree" to "strongly agree" with a neutral mean ("neither nor").		

Prior to exposing respondents to the questions described above, an introductory text was presented that described the proposed behavior-based automobile insurance model. Technical details were deliberately not disclosed in order to avoid the situation where respondents evaluate the perceived technical feasibility of the approach. Instead, the description was formulated to suggest that the technical implementation will be available without revealing any details. The following text has been presented at the beginning of the survey:

Please imagine you want to close a new automobile insurance contract. A well-known automobile insurance company is offering you a novel contract, under which your insurance rates will depend on your driving behavior. Your garage will install a device in your vehicle at no charge. This device will observe your driving behavior and send this information to your insurance company. The installation will be carried out at a location in your vehicle that is not visible.

Your insurance rate will depend on your driving behavior. If you drive carefully and cautiously, you will pay less than under a classical contract. If you drive however aggressively and incautiously, you will pay more than under a classical contract. To evaluate your driving style, various measurement data are sent from your vehicle to the insurance company, which will allow for a reliable estimation of your accident risk.

On a special web site and on your smartphone, you can always see how your driving behavior is rated, and how this affects your insurance rate.

IV.4.2 Sample

The data for the present study were collected via an online survey in November 2011. Study participants were recruited by a market research agency. Financial compensation was granted for a completely filled in questionnaire. The survey took about 12 minutes to complete. The sample is representative for the German population above the age of 18 years and in possession of a driving license. People without a driving license were excluded from the survey, as we can assume that they lack the experience required to accurately evaluate the proposed automobile insurance model. The resulting sample comprised 315 subjects, which were well balanced across gender and age (Table 48). To determine the minimum sample size, the software tool G*Power was used (Faul et al. 2009). Since the maximum number of predictors is eight, a sample size of 160 is required to detect even small effect sizes above 0.15 with a statistical power of 0.95 at $\alpha=0.05$. For the full sample, this requirement is largely fulfilled with n=315. For the analysis of the moderating effect of gender however, the sample size for women is only 156. Power analysis shows that, in this case, effects above an effect size of 0.154 can be detected, which is still acceptable.

Table 48: Sample Demographics

Characteristic	Category	Frequency	in %
Gender	Female	156	49.5 %
	Male	159	50.5%
Age	18 - 27	58	18.4%
	28 - 37	63	20.0%
	38 - 47	70	22.2%
	48 - 57	64	20.3%
	58 - 67	45	14.3%
	>68	15	4.8%
Sample size n = 315			

IV.5 Data Analysis

IV.5.1 Analysis Methodology

The purpose of this study is to analyze how consumers perceive the proposed automobile insurance model, and to predict whether such an insurance model may induce a change in driving behavior. Although the proposed research model is rooted in theoretical results of technology acceptance research, new constructs and relationships were added to account for the specific aspects of a novel insurance business model. Therefore, the analysis approach is not strictly confirmatory as we intend to explore to which degree certain relationships may influence consumer behavior.

PLS-based SEM approaches are superior to covariance-based analysis procedures for exploratory studies that are not completely grounded in existing theories (Ainuddin et al. 2007; Henseler et al. 2009; Holzmüller and Kasper 1991; Jöreskog and Wold 1982). We will therefore apply PLS to evaluate the proposed research model, although sample size and distributional assumptions would also justify the application of a covariance-based approach (see section II.5.1.2 for a detailed discussion of PLS-based structural equation modeling). For estimating the research model, the software SmartPLS version 2.0M3 (Ringle et al. 2005) was used.

Our data analysis follows the same approach as described in section II.5.1 (Figure 15). First, the measurement model is validated by means of confirmatory factor analysis. Then the structural model is assessed by estimating model parameters, path significances, effects sizes and the predictive relevance of the model. In a next step, the hypothesized moderator effects are tested by multi-group comparison for the categorical variable *Gender* and by applying the product-indicator approach for the continuous variables *Cost Sensitivity* and *Driving Style*. Finally, we conduct an exploratory data analysis to investigate further aspects of the proposed insurance model. For this purpose, we descriptively analyze additional questions we raised in our survey.

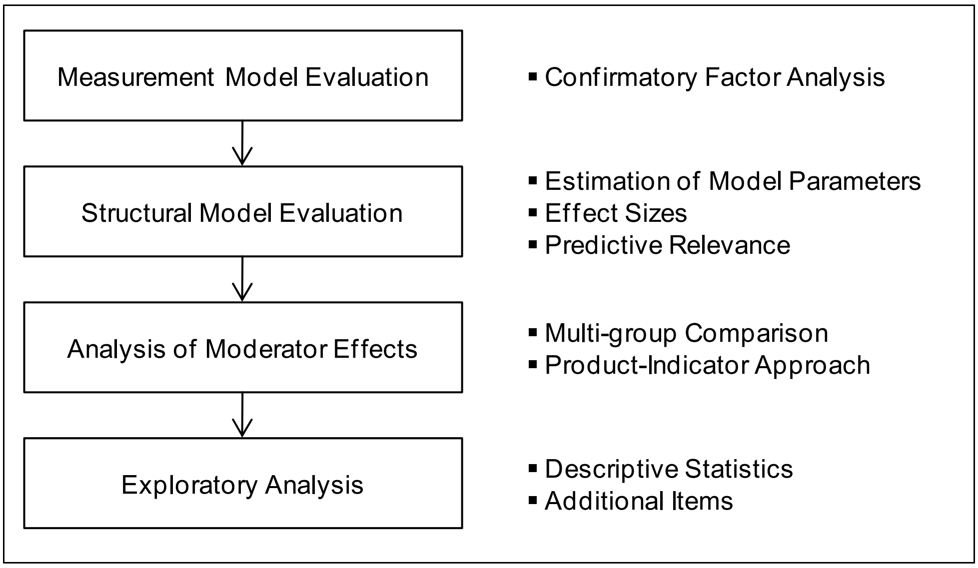


Figure 15: Data Analysis Process

IV.5.2 Measurement Model

To assess the validity of the measurement model, a confirmatory factor analysis was conducted (Table 49). With the exception of PP3, all items show factor loadings well above the usually recommended threshold of 0.707, which indicates that more than half of the item variance is captured by the latent construct (Chin 1998a; Gefen et al. 2000). PP3 has an indicator loading of only 0.67 and is therefore slightly below the threshold. However, since all construct level criteria are fulfilled, and as PP3 covers an important aspect from a content perspective, we decided to keep this item in the model.

Internal consistency of the constructs is assessed by Cronbach's α and Composite Reliability. Cronbach's α values all exceed the usually applied threshold of 0.7 (Nunnally and Bernstein 1994), and most constructs have α values above 0.8. Composite reliability values, which are a more appropriate measure for internal consistency when PLS is applied, are also well above the recommended threshold of 0.7 (Nunnally and Bernstein 1994). Convergent validity is assessed by the means of the Average Variance Extracted (AVE), which should lie above the threshold of 0.5 (Fornell and Larcker 1981). For our measurement model, all AVE values are larger than 0.5.

Table 49: Validation of the Measurement Model

Construct	Item	Loading	Mean	Std. Dev.	LV Score	α	CR	AVE
Attitude (AT)	AT1	0.91	5.02	1.70	4.85	0.94	0.96	0.81
	AT2	0.90	4.84	1.46				
	AT3	0.85	4.76	1.46				
	AT4	0.93	4.92	1.68				
	AT5	0.92	4.69	1.73				
Behavioral Intention (BI)	BI1	0.94	4.46	1.59	4.64	0.95	0.97	0.87
	BI2	0.93	4.87	1.60				
	BI3	0.94	4.48	1.63				
	BI4	0.92	4.76	1.50				
Behavioral Intention to Change (BIC)	BIC1	0.91	4.45	1.58	4.50	0.93	0.96	0.88
	BIC2	0.96	4.55	1.55				
	BIC3	0.95	4.49	1.48				
Trust (TR)	TR1	0.96	4.34	1.48	4.57	0.86	0.91	0.77
	TR2	0.96	4.43	1.48				
	TR3	0.69	5.00	1.20				
Effort Expectancy (EE)	EE1	0.73	5.50	1.23	4.94	0.76	0.85	0.58
	EE2	0.81	4.97	1.39				
	EE3*	0.78	4.72	1.54				
	EE4*	0.72	4.44	1.59				
Performance Expectancy (PE)	PE1	0.92	4.43	1.52	4.55	0.90	0.94	0.83
	PE2	0.86	4.74	1.46				
	PE3	0.94	4.47	1.49				
Perceived Enjoyment (PJ)	PJ1*	0.88	4.02	1.61	4.39	0.88	0.92	0.74
	PJ2*	0.88	4.29	1.51				
	PJ3*	0.90	4.18	1.58				
	PJ4	0.77	5.11	1.43				
Perceived Privacy (PP)	PP1	0.86	4.68	1.56	4.09	0.74	0.86	0.67
	PP2	0.90	4.10	1.60				
	PP3*	0.67	3.28	1.59				
Relative Advantage (RA)	RA1	0.90	4.78	1.49	4.87	0.90	0.93	0.77
	RA2	0.88	4.92	1.40				
	RA3	0.90	4.91	1.50				
	RA4	0.82	4.86	1.55				
Social Influence (SI)	SI1	0.95	4.39	1.42	4.44	0.94	0.96	0.90
	SI2	0.95	4.33	1.40				
	SI3	0.95	4.60	1.32				
Cost Sensitivity (CS)	CS1	0.79	5.52	1.11	5.49	0.78	0.87	0.69
	CS2	0.87	5.53	1.18				
	CS3	0.83	5.40	1.33				
Driving Style (DS)	DS1	0.75	5.64	1.06	5.41	0.76	0.83	0.62
	DS2	0.69	5.73	0.96				
	DS3	0.90	4.83	1.27				

Items marked with (*) were measured on a reverse scale. SD: standard deviation; LV Score: latent variable score; α: Cronbach's α; CR: Composite Reliability; AVE: average variance extracted

The Fornell/Larcker criterion is applied to assess the discriminant validity of the measurement model, which describes the degree to which measures of different concepts differ. The Fornell/Larcker criterion postulates that the squared root of the AVE of each construct is larger than the correlations of this construct with all other constructs (Fornell and Larcker 1981). Table 50 shows that this criterion is fulfilled for all constructs so that sufficient discriminant validity can be confirmed.

Table 50: Discriminant Validity

Constructs												
	AT	**BI**	**BIC**	**TR**	**EE**	**PE**	**PJ**	**PP**	**RA**	**SI**	**CS**	**DS**
AT	**0.90**											
BI	0.80	**0.93**										
BIC	0.58	0.68	**0.94**									
TR	0.65	0.76	0.57	**0.88**								
EE	0.48	0.52	0.23	0.46	**0.76**							
PE	0.76	0.83	0.62	0.77	0.45	**0.91**						
PJ	0.59	0.72	0.42	0.56	0.60	0.65	**0.86**					
PP	0.62	0.71	0.48	0.77	0.37	0.68	0.51	**0.82**				
RA	0.75	0.84	0.63	0.67	0.53	0.74	0.64	0.61	**0.88**			
SI	0.70	0.65	0.51	0.55	0.30	0.60	0.42	0.51	0.64	**0.95**		
CS	0.32	0.32	0.32	0.29	0.14	0.30	0.19	0.23	0.28	0.23	**0.83**	
DS	0.27	0.30	0.10	0.26	0.29	0.29	0.33	0.17	0.30	0.16	0.37	**0.79**
Note: The diagonal reports the square roots of the average variance extracted (AVE).												

We conclude from the confirmatory factor analysis that the measurement model sufficiently fulfills the applied reliability and validity criteria. Consequently, we will take this measurement model without further modifications as the basis for evaluating the structural model.

IV.5.3 Structural Model

To assess the explanatory power of the proposed research model, we firstly inspect the R^2 values for the dependent variables, which describe the percentage of variance explained by their antecedents. R^2 values of 0.71 and 0.85 for AT and BI can be regarded as substantial, and an R^2 of 0.46 for BIC indicates a moderate level of explained variance (Chin 1998a). Overall, R^2 values for the three dependent variables indicate a satisfactory and substantive level of explanatory power for our model.

To test the hypotheses for the main effects (i.e. without moderators), we apply bootstrapping as a re-sampling technique to obtain pseudo-t-values for evaluating the significance level of the path weights (Chin 1998a). Figure 16 shows the path weights and significance levels for the main effects, as obtained from the PLS estimation with bootstrapping.

PE has a strong and significant effect on AT and BI. Consequently, we accept hypotheses H1 and H2. The effect of EE on AT and BI is not significant (H3 and H4 are rejected). SI has a strong and significant effect on AT, but not on BI (H5 rejected, H6 accepted). The effect of PJ on BI is significant, while PJ does not significantly influence AT (H7 accepted, H8 rejected). RA has a strong and significant effect on both AT and BI (H9 and H10 accepted). PP has a weak, but significant effect on BI, yet not on AT (H11 accepted, H12 rejected). TR has a weak, but significant effect on BI (H13 accepted). Attitude has a moderate and significant effect on BI (H14 accepted). Finally, BI has a very strong and significant effect on BIC (H15 accepted).

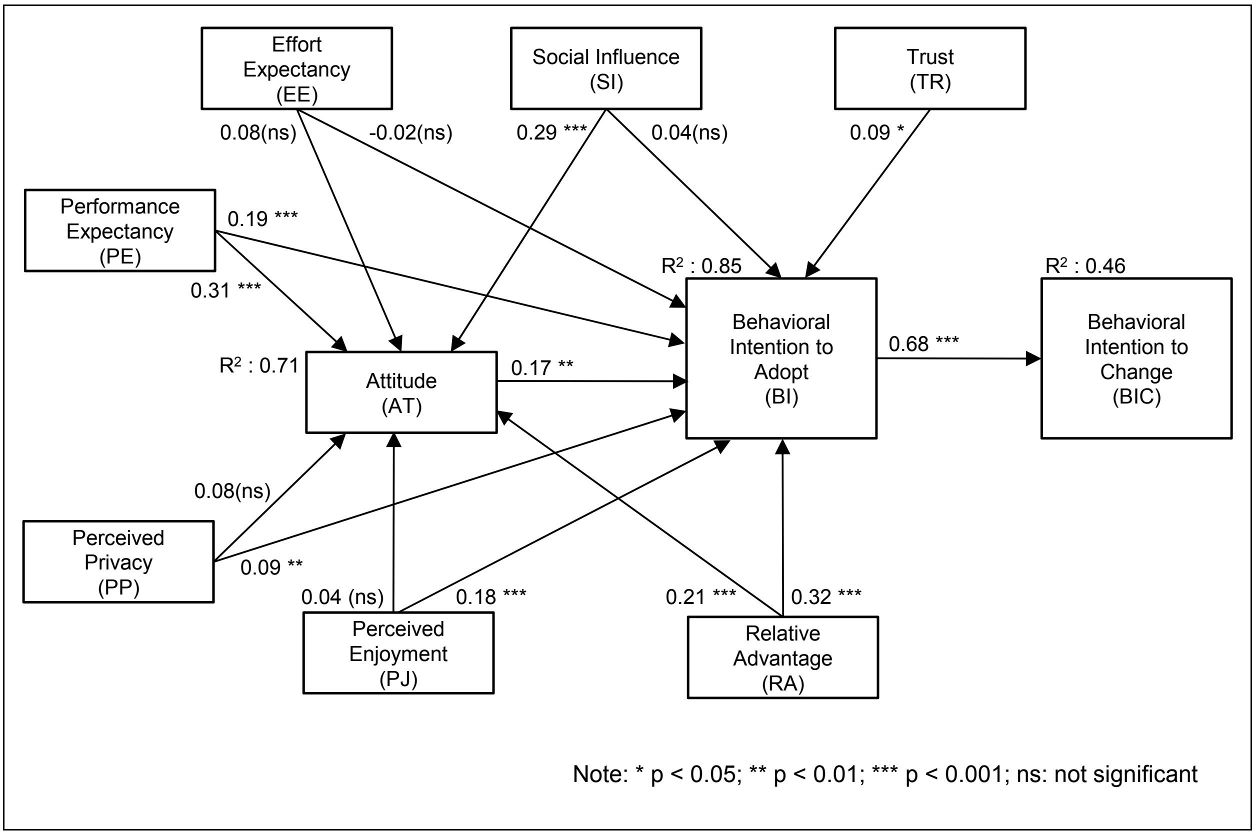


Figure 16: Results for the Structural Model

Effect sizes (Cohen's f^2) were calculated to assess the explanatory power of each individual exogenous variable with regards to *Attitude* and *Behavioral Intention to Adopt*

(Table 51). Values of 0.02, 0.15, and 0.35 can be regarded as "small", "medium", and "large" effect sizes (Chin 1998a; Cohen 1988).

We can see that all significant paths in the structural model have at least small but substantial effect sizes, whereas insignificant paths do not have a substantial effect size. The largest effect size was found for the RA-BI relationship, followed by SI-AT, PE-AT, and PJ-BI. All other effect sizes are substantially smaller. Relationships that show insubstantial effect sizes were also found to have insignificant path weights.

Table 51: Effect Sizes

Construct	Effect Size f^2 (AT)	Interpretation	Effect Size f^2 (BI)	Interpretation
Trust (TR)	-	-	0.02	small
Effort Expectancy (EE)	0.01	not substantial	0.00	not substantial
Performance Expectancy (PE)	0.11	small	0.06	small
Perceived Enjoyment (PJ)	0.00	not substantial	0.10	small
Perceived Privacy (PP)	0.01	not substantial	0.02	small
Relative Advantage (RA)	0.05	small	0.21	medium
Social Influence (SI)	0.16	medium	0.01	not substantial

To assess the predictive relevance of the model, we apply Stone-Geisser's Q^2 criterion for the three endogenous variables AT, BI, and BIC. The values of 0.57, 0.73, and 0.40, which were obtained by applying the blindfolding procedure implemented in SmartPLS, are well above zero, which confirms the predictive relevance of the model (Chin 1998a; Tenenhaus et al. 2005).

To obtain a measure for the relative predictive relevance of each variable, we subsequently eliminate each exogenous variable from the model and calculate the resulting Q^2 value for the reduced model. Analogous to Cohen's f^2, we obtain q^2 as a measure for the relative contribution of each exogenous variable to the overall predictive relevance Q^2 of the model by calculating the relative change of Q^2. As for f^2, values of 0.02, 0.15, and 0.35 represent "small", "medium", and "large" levels of predictive relevance respectively (Henseler et al. 2009). Table 52 shows that each individual variable has only a small or insubstantial level of predictive relevance. Comparably large values were obtained for the relationships RA-BI, SI-AT, PE-AT, and PJ-BI. We can further see that the relationships TR-BI and PI-BI have no substantial predictive relevance, although their path weights are significant. All other relationships for which the predictive relevance is not substantial were also found to have insignificant path weights.

Table 52: Predictive Relevance

Construct	Predictive Relevance q^2 (AT)	Interpretation	Predictive Relevance q^2 (BI)	Interpretation
Trust (TR)	-	-	0.01	not substantial
Effort Expectancy (EE)	0.01	not substantial	0.00	not substantial
Performance Expectancy (PE)	0.06	small	0.03	small
Perceived Enjoyment (PJ)	0.00	not substantial	0.05	small
Perceived Privacy (PP)	0.00	not substantial	0.01	not substantial
Relative Advantage (RA)	0.03	small	0.10	small
Social Influence	0.08	small	0.00	not substantial

IV.5.4 Moderator Effects

For evaluating the moderating effect of the categorical variable *Gender*, we conducted a PLS estimation for the two data subsets for male and female respondents. An adapted t-test is applied to test whether path weights that are relevant to our hypotheses significantly differ between the two subsets (Ahuja and Thatcher 2005; Keil et al. 2000; Rai and Keil 2008; Venkatesh and Morris 2000). Table 53 shows that all hypothesized moderation effects by gender are significant. Splitting the sample by gender, the SI-BI relationship becomes insignificant for men, whereas it becomes substantially stronger for women, compared to the whole sample. Vice-versa, the AT-BI relationship becomes much stronger for men, but insignificant for women. The influence of PJ and RA on BI is higher for men than for women. Consequently, we accept all gender-related hypotheses H16a, H16b, H16c, and H16d.

Table 53: Moderating Effect of Gender

Moderator		R^2		Path Coefficients			
		BI	AT	SI → BI	AT → BI	PJ → BI	RA → BI
None		0.851	0.712	0.04 ns	0.17 **	0.18 ***	0.32 ***
Gender	Male	0.850	0.721	-0.02 ns	0.28 **	0.16 ***	0.27 ***
	Female	0.868	0.741	0.11 ***	0.05 ns	0.25 ***	0.36 ***
	T-Test			***	***	*	*
Note: * p < 0.05; ** p < 0.01; *** p < 0.001; ns: not significant							

For testing the moderating effects of the continuous variables *Cost Sensitivity* and *Driving Style*, we applied the product-indicator approach as described by Chin, Marcolin, and Newsted (2003). A t-test reveals the significance levels of both the direct effect of the moderator variable on the dependent variable BIC, and its interaction effect on the relationship between BI and BIC. Additionally, an F-test is applied to decide whether the explained variance R^2 of BIC is significantly improved by adding the moderator variable.

Table 54 shows the results of this moderator analysis. *Cost Sensitivity* has a significant direct effect on BIC and significantly moderates the BI-BIC relationship. Persons who are more cost sensitive show a higher willingness to change their driving behavior under a behavior-based insurance contract. As the path weights and the improvement of R^2 are significant, we accept hypotheses H17a and H17b.

Driving Style has a significant direct effect on BIC, but does not significantly moderate the BI-BIC relationship. The direct effect of DS on BIC is negative, which means that drivers with a cautious driving style have a lower tendency to change their driving behavior under a behavior-based insurance contract. The change of the explained variance is significant, so we accept H18a (direct effect), but reject H18b (moderator effect).

Table 54: Moderating Effects of Cost Sensitivity and Driving Style

Moderator		**R^2 (BIC)**	**Path Coefficients** BI → BIC
None		0.4586	0.68 ***
Cost Sensitivity (CS)	Direct Effect	0.4860	0.13 **
	Interaction	-	0.13 *
	F-Test	16.63 ***	-
Driving Style (DS)	Direct Effect	0.4697	-0.11 *
	Interaction	-	0.01 ns
	F-Test	6.53*	-
Note: * p < 0.05; ** p < 0.01; *** p < 0.001			

IV.5.5 Exploratory Analysis

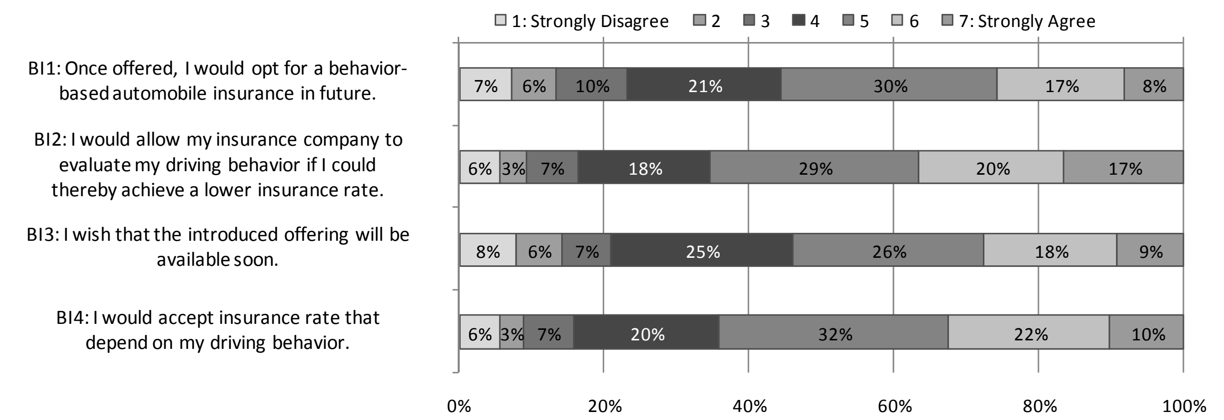
In the structural equation analysis presented in the previous sections, we investigated how different factors influence the intention to adopt behavior-based automobile insurance and the intention to change one's driving behavior in the presence of such a contract. In this section we explore further details regarding the perception of behavior-based insurance models. First we analyze the BI and BIC constructs as the central outcome variables of the research model. Then we describe the results obtained from additional questions raised in the survey, in order to draw practical conclusions for a market introduction of behavior-based automobile insurance models.

IV.5.5.1 ***Behavioral Intention to Adopt***

Questions related to the *Behavioral Intention to Adopt* construct asked survey participants about their willingness to accept behavior-based automobile insurance. Between 53% (BI3) and 64% (BI4) of the answers were positive, whereas only between 16% (BI2, BI4) and 23% (BI1) of the answers were negative (Table 55). For all four items, the most often answer was "I rather agree", i.e. a value of "5" on the Likert scale. T-tests reveal that the mean value of all items is significantly larger than the neutral value of 4. These results indicate that there is a tendency among the respondents of the present survey to be willing to adopt a behavior-based insurance contract.

Table 55: Behavioral Intention to Adopt

Behavioral Intention to Adopt (BI)



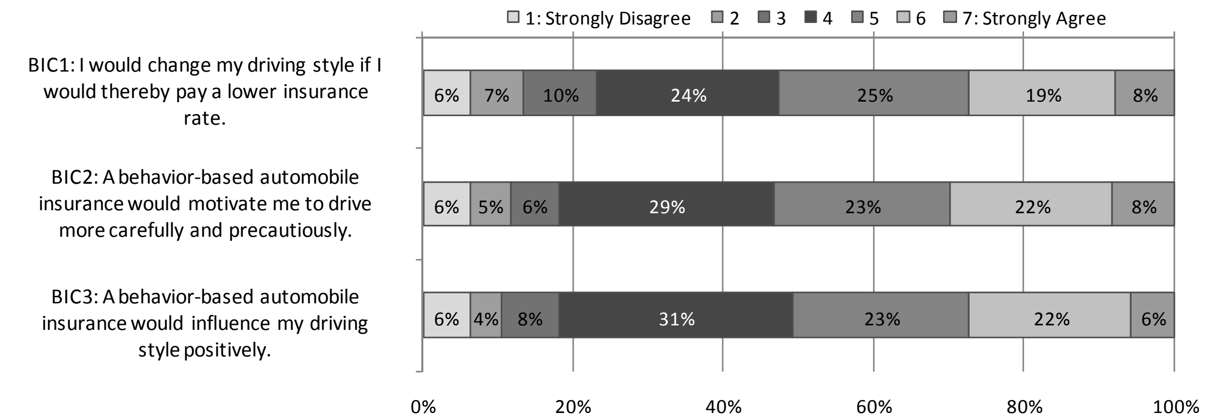
Item	Mode	Mean	Std. Dev.	Deviation from scale mean	Sig. (2-tailed)
BI1	5	4.46	1.59	0.46	.000 ***
BI2	5	4.87	1.60	0.87	.000 ***
BI3	5	4.48	1.63	0.48	.000 ***
BI4	5	4.76	1.50	0.76	.000 ***

Note: * p < 0.05; ** p < 0.01; *** p < 0.001

IV.5.5.2 ***Behavioral Intention to Change***

The purpose of the *Behavioral Intention to Change* construct is to explore whether people would change their driving style under a behavior-based automobile insurance contract. Between 51% (BIC3) and 53% (BIC2) of the respondents agreed that they would change their driving style, whereas between 17% (BIC2) and 22% (BIC1) of the respondents disagreed (Table 56). The means of the three items are all significantly larger than the neutral value of four. For BIC1, respondents most often answered that they agreed. For BIC2 and BIC3, the most frequent answer was the neutral value of four on the Likert scale ("neither nor"). We conclude that the study participants expect that they would change their driving behavior under an insurance contract that grants monetary rewards for careful driving. Although the result may not allow for the conclusion that such a behavioral change will indeed occur, it demonstrates at least that the majority of our respondents is basically willing to adapt their driving behavior in exchange for a monetary benefit.

Table 56: Behavioral Intention to Change



Item	Mode	Mean	Std. Dev.	Deviation from scale mean	Sig. (2-tailed)
BIC1	5	4.45	1.58	0.45	.000 ***
BIC2	4	4.55	1.55	0.55	.000 ***
BIC3	4	4.49	1.48	0.49	.000 ***
Note: * p < 0.05; ** p < 0.01; *** p < 0.001					

IV.5.5.3 ***Expected Saving Potential***

Behavior-based automobile insurance models imply a trade-off between the chance of getting a financial reward for adopting an appropriate driving style on the one hand and a potentially negative impact on driving pleasure as well as privacy risks on the other hand. Consequently, it is a key question for a successful market introduction of behavior-based insurance models, how large the achievable saving potential must be so that consumers accept the negative side effects. To explore how the offered saving potential impacts market acceptance, we raised the following question:

Imagine you want to close an automobile insurance contract, which normally costs 700 EUR per year. How much savings must be achievable so that you decide for a behavior-based automobile insurance contract?

The results illustrated in Figure 17 show that the most frequent answer was "more than 200 EUR". A cumulated proportion of 50% of the respondents stated that they would decide for the proposed insurance model under this condition. If the saving potential is more than 150 EUR but less than 200 EUR, only 25% of the respondents are interested in behavior-based insurances. 16% expect even more than 400 EUR. Summarizing these results, 200 EUR seems to be a threshold for many respondents, above which they would perceive the trade-off between positive and negative effects of the proposed offering as personally advantageous.

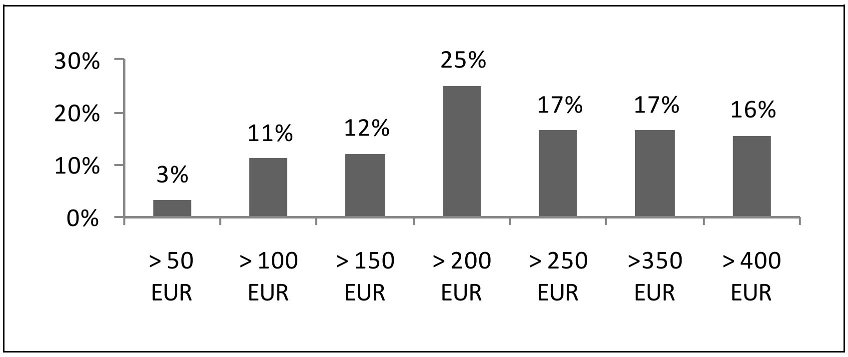
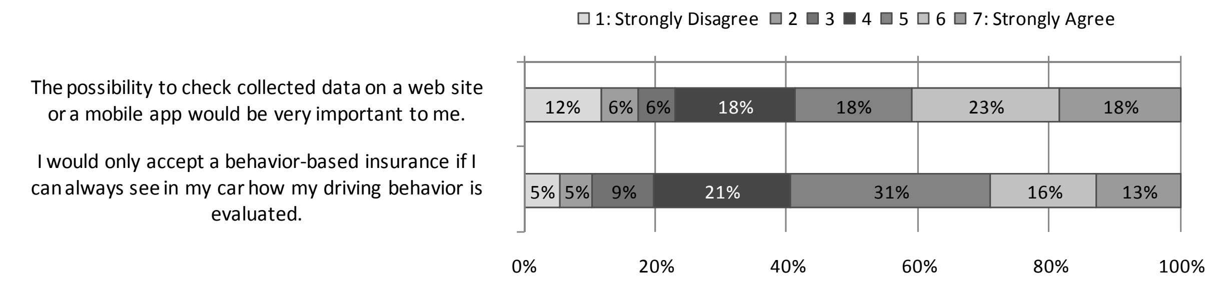


Figure 17: Expected Saving Potential

IV.5.5.4 ***Feedback on Driving Behavior***

In the introductory text of the survey, we stated that the proposed insurance model would come along with a web site and a mobile phone application, with which customers can control how their driving behavior is evaluated and how this evaluation impacts their insurance rate. With regard to this feature, we first asked our respondents how important it is for them to get such an online report. Second, we asked whether real-time feedback in the vehicle is regarded as a prerequisite for accepting a behavior-based insurance model. The mean values of the responses indicate that both features are similarly important (Table 57). The mode values of 6 for the first and 5 for the second question indicate that feedback via an online service might be more important to consumers than real-time feedback in the vehicle. At the same time, 24% of the respondents rated online feedback as rather unimportant, 19% disagreed with the notion that they would only accept a behavior-based insurance model if real-time feedback is provided.

Table 57: Feedback on Driving Behavior



Item	Mode	Mean	Std. Dev.	Deviation from scale mean	Sig. (2-tailed)
Web Site / App	6	4.66	1.91	0.66	.000 ***
Real-time feedback	5	4.65	1.57	0.65	.000 ***
Note: * p < 0.05; ** p < 0.01; *** p < 0.001					

IV.5.5.5 ***Data Provisioning***

At the beginning of the survey, we gave only a rough explanation about the technical implementation used to evaluate one's driving behavior. We stated: *"This device will observe your driving behavior and send this data to your insurance company"*. We deliberately did not indicate which data will be collected and how it is evaluated. At the end of the survey, we asked our respondents which of the following information they would *not* be willing to share with their insurance company (Figure 18).

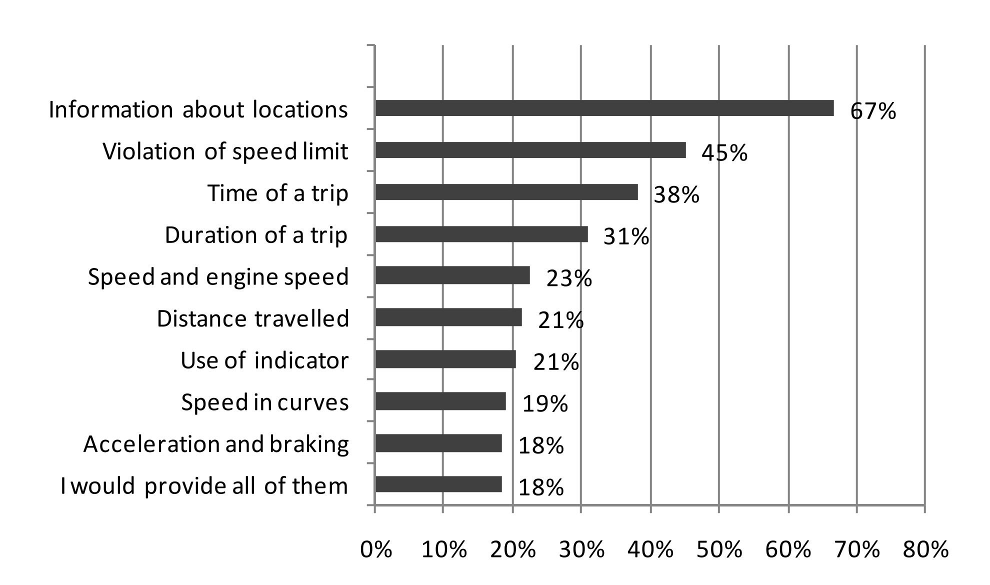


Figure 18: Resistance to Provide Data

Information about locations is regarded as most critical (67%), followed by *violation of speed limits* (45%) and *time of a trip* (38%). Comparably low resistance was stated for *acceleration and braking* (18%), *speed in curves* (19%) and *speed / engine speed*

(23%), which are important factors for evaluating driving behavior. Only 18% of the respondents would be willing to share all of the mentioned data.

IV.6 Discussion

IV.6.1 Theoretical Implications

The present study contributes to the body of work in the field of technology acceptance research by decomposing and adapting existing technology acceptance models such as TAM and UTAUT to evaluate the perception of a behavior-based automobile insurance model. Established technology acceptance models like TAM (Davis 1989) and UTAUT (Venkatesh et al. 2003) were developed with the goal to obtain a parsimonious explanation model that explains adoption decisions by a very limited number of influencing factors. Whilst being appropriate for the analysis of information systems acceptance in a professional setting, subsequent research has shown that complex technology-based business models in the private domain require more fine-grained analysis models to capture the different dimensions that explain consumer intentions and purchase decisions. For example, decomposed and extended analysis models on the basis of TAM and UTAUT were developed with regard to the use of web sites (Atkinson and Kydd 1997; Heijden 2004; Moon and Kim 2001), private use of computers (Brown and Venkatesh 2005), electronic commerce (Chellapa and Sin 2005; Dinev and Hart 2006; Eastin 2002; Gefen et al. 2003; Mathwick et al. 2001; Mukherjee and Nath 2007; Pavlou and Gefen 2004; Roca et al. 2009; Sheng et al. 2008), mobile services (Hong et al. 2008; Nysveen et al. 2005), mobile banking (Luarn and Lin 2005; Shin 2009), and mobile ticketing (Mallat et al. 2006).

Typically, these models extend TAM and UTAUT by adding constructs that refer to *Trust* (Chellapa and Sin 2005; Dinev and Hart 2006; Gefen et al. 2003; McKnight et al. 2002; Pavlou and Gefen 2004; Roca et al. 2009), *Privacy* (Chellapa and Sin 2005; Dinev and Hart 2006; Mukherjee and Nath 2007; Roca et al. 2009), *Relative Advantage* and *Financial Benefit* (Eastin 2002; Hong et al. 2008; Luarn and Lin 2005; Mallat et al. 2006; Moore and Benbasat 1991), *Perceived Enjoyment and Playfulness* (Atkinson and Kydd 1997; Brown and Venkatesh 2005; Dabholkar and Bagozzi 2002; Heijden 2004; Hong et al. 2008; Mallat et al. 2006; Moon and Kim 2001; Nysveen et al. 2005), *Self-Efficacy* (Eastin 2002; Luarn and Lin 2005), and *Attitude* (Mallat et al. 2006; Moon and Kim 2001; Nysveen et al. 2005; Shin 2009). As a synthesis of the various extensions of TAM and UTAUT, which are present in technology acceptance research, we developed a research model that is specific to the analysis of behavior-based automobile insurance.

Results obtained from confirmatory factor analysis demonstrate that the adapted measurement scales achieve a sufficient level of reliability and validity. The structural model explains 85% of the variance of the central outcome variable *Behavioral Intention to Adopt*. For *Attitude* and *Behavioral Intention to Change,* we obtained R^2 values of 0.71 and 0.46 respectively.

The *Behavioral Intention to Change* construct was newly developed for this study to explore to which degree the intention to adopt the proposed insurance model influences the willingness to change driving behavior. A path weight of 0.68 indicates that behavior-based automobile insurances may have a positive impact on driving style. Two context-specific variables have been assumed to moderate the relationship between *Behavioral Intention to Adopt* and *Behavioral Intention to Change*.

The direct effect of *Driving Style* has shown that people who are adopting a reckless driving style are more willing to change than people who already drive carefully. The hypothesized moderating effect of *Driving Style* on the BI-BIC relationship could not be confirmed. The BI-BIC relationship is moderated by *Cost Sensitivity*. The more cost sensitive a person is, the more is he willing to change driving behavior under a behavior-based insurance contract. Additionally, *Cost Sensitivity* has a significant direct effect on the *Behavioral Intention to Change*.

From a theoretical perspective, introducing the BIC construct was a first attempt to extend the technology acceptance model to better meet the goals of the persuasive technology research domain. Evaluating persuasive technologies, researchers are not only interested in predicting the degree to which a persuasive approach may be accepted by potential users, but also to assess the degree to which it may achieve its goal to induce a sustainable change in behavior. The BIC construct was developed to measure the willingness to change one's behavior under the condition of the proposed insurance model. By evaluating the relationship between BI and BIC, we could infer on the degree to which the willingness to adopt a persuasive technology or business model influences the willingness for behavioral change.

Relative Advantage is the most salient factor to influence *Behavioral Intention to Adopt* with a path weight of 0.32, and has a strong influence of 0.21 on *Attitude*. To a lesser but still strong degree, *Performance Expectancy* (0.19) and *Perceived Enjoyment* (0.18) influence *Behavioral Intention to Adopt.* Evidently, potential adopters primarily consider their chance of getting an appropriate financial reward under a behavior-based insurance contract. The expected financial advantage is in contrast to a potential loss of driving pleasure and the risk of error-prone technology. Our results have shown that adoption decisions are primarily determined by balancing out the trade-off between these three factors.

Trust and *Perceived Privacy* play only a minor, but nonetheless significant role for predicting the *Behavioral Intention to Adopt*. The majority of our respondents have enough trust in their insurance company to perceive the proposed insurance model as fair. They also trust that data would not be exploited in an inappropriate way. The risk of privacy issues due to technical insufficiencies is also not regarded as very critical. Although trust and privacy considerations play some minor role, the degree to which they influence the adoption decision is surprisingly low.

The influence of *Effort Expectancy* on *Behavioral Intention to Adopt* and on *Attitude* is not significant. The majority of our respondents expect to be capable of dealing with the increased complexity of behavior-based insurance contracts so that effort considerations are negligible for their adoption decision.

Social Influence had no significant impact on the *Behavioral Intention to Adopt*, but a strong effect on *Attitude*. Evidently, people tend to align their general attitude in accordance with the social norms imposed by their environment. In contrast, the adoption decision is made on the basis of self-centered criteria, which are primarily the expected financial benefit, technical feasibility and consequences for their personal enjoyment.

The general A*ttitude* towards the proposed insurance model is strongly influencing the adoption decision. Furthermore, *Attitude* mediates the influence of *Performance Expectancy*, *Social Influence*, and *Relative Advantage* on BI, which explain 71% of the variance of AT. The strong influence of RA shows that people cannot shape their opinion about the proposed insurance model without taking into account their personal financial benefit. However, the influence of RA is relatively small compared to SI and PE, which are the primary factors in influencing the general attitude towards behavior-based automobile insurances.

Gender was found to moderate the effect of *Social Influence, Attitude, Perceived Enjoyment,* and *Relative Advantage* on the *Behavioral Intention to Adopt*. The SI-BI relationship was found to be significant only for women, whereas it is insignificant for men and the whole sample. This supports the evidence from prior studies that women have a stronger tendency than men to align their decisions with social norms and expectations (Venkatesh et al. 2003, 2000; Venkatesh and Morris 2000). Vice-versa, the

AT-BI relationship is only significant for men but not for women - a result that is again in line with prior studies on gender differences (Venkatesh et al. 2000; Venkatesh and Morris 2000). Analysis results for the SI-BI and AT-BI relationship support the prior finding from the field of psychological research, which states that men are generally more task-oriented, have higher levels of self-esteem, and base their decisions primarily on objective criteria, whereas women have a stronger tendency to comply with social norms which they believe would elicit a favorable reaction from others (Bem 1981; Carli 2001; Carlson 1971; Eagly 1978; Hoffman 1972; Lundeberg et al. 1994; Venkatesh et al. 2000).

The extent to which *Perceived Enjoyment* and *Relative Advantage* influence the *Behavioral Intention to Adopt* has been found to be larger for women than for men. Considering previous research results, the opposite effect could have been expected as men have consistently been found more than women to base their decisions on objective and logical arguments. In our specific context however, men can be expected to benefit less than women from a behavior-based insurance model and to worry about a potentially negative influence on driving pleasure due to their stronger disposition for a risky and reckless driving style. Although contradicting to previous findings, our results therefore reflect a rational consumer decision. They further demonstrate that gender differences are not generally transferrable to any application domain. One's general attitude towards the specific technology or business model, perceptions and behavioral patterns evidently influence the role of gender differences.

From a theoretical perspective, our results demonstrate that context-specific extensions of technology acceptance models can produce relevant insights with regards to the relationship between technology adoption and its influence on behavioral patterns. The decomposition of more parsimonious models helps to explain how different factors interact when a consumer is confronted with an adoption decision. The analysis of effect sizes has shown that none of the investigated predictors has a dominant role. Instead, all significant factors make a rather small to medium contribution to the explanation of the outcome variable, which accumulate to a comparably high explained variance of 85% for the *Behavioral Intention to Adopt.* Further research is required to investigate whether the extensions we introduced may be a fruitful approach for evaluating the effectiveness of persuasive technologies and persuasive business models in other application domains.

IV.6.2 Practical Implications

The exploratory approach taken in this study reveals a number of practical implications for the automobile insurance industry. Most important, more than half of our respondents stated that they would, in principle, be willing to close a behavior-based automobile insurance contract in exchange for a monetary reward for careful driving. The magnitude of this monetary compensation turned out to be the major determinant for an adoption decision. Assuming costs of 700 EUR for a conventional insurance contract, 50% of our respondents indicated their willingness to accept the proposed insurance model under the condition that a saving potential between 200 and 250 EUR is achievable. In contrast to that, the acceptance rate goes back to 14% if savings between 100 and 150 EUR can be achieved. The comparably high latent variable score of 4.85 for *Attitude* indicates a positive general attitude towards the proposed insurance product, but *Behavioral Intention to Adopt* primarily depends on the expected financial advantage over a conventional product.

Performance Expectancy, i.e. the believe that the underlying technology will allow for an appropriate evaluation of driving behavior, is another important influencing factor for the general attitude towards the product and the willingness to adopt it. From a practical perspective, emphasis should be put on explaining the technical implementation to convince critical customers that the technology will be capable of determining accurate risk profiles. Evidently, there is a strong desire for transparency with regard to the risk assessment. A majority of almost 60% of our respondents stated that it would be important for them to get online or real-time feedback on their driving style and its consequences for their insurance premium. Without such detailed information, only a minority of about 20% would be willing to accept the proposed insurance model. Therefore a strong focus should be put on offering online feedback via a website and real-time feedback, for example via a smartphone application or a dedicated device, to fulfill the desire for information transparency, and to mitigate remaining doubts about the technical feasibility of accurately evaluating driving behavior.

In contrast to our expectations, trust and privacy considerations play only a minor role with regard to an adoption decision. The respective latent variable scores indicate that our respondents are confident that the collected data would not be abused, and that the proposed insurance product would be a trustworthy offering. The influence of both aspects on *Behavioral Intention to Adopt* is only weak. We therefore conclude that these aspects should rather be regarded as a "hygienic factor". Special effort to foster trust might not be necessary, but every effort must be made that the currently sufficient level of credibility is not affected by negative experiences. Examples from other domains (e.g. RFID, social networks) have shown that negative information about privacy issues can dominate the discussion about a product despite primarily positive consumer attitudes at the launch of a novel technology.

The strong effect of *Social Influence* on *Attitude*, which significantly predicts the *Behavioral Intention to Adopt*, indicates that a positive public opinion may be important for a successful commercialization of the proposed insurance model. Textual comments we received from a number of respondents show that two expectations may contribute to a positive image. One group appreciated that the proposed insurance model implies a punishment for drivers that endanger other traffic participants. A second group stated that they would perceive insurance premiums fairer as they would not be accounted for careless drivers. These perceptions may be taken as a starting point to develop a marketing strategy to foster a positive product image.

Effort Expectancy turned out to be non-influential for a potential adoption decision. Although behavior-based insurance contracts severely increase the complexity of the insurance product, our respondents did not feel overburdened by controlling their driving style or getting used to changing insurance premiums.

Results for the *Perceived Enjoyment* construct show that people, who do not perceive the proposed insurance model as a restriction to their driving pleasure, have a significantly stronger intention to adopt. Assuming that careful drivers feel less affected, the proposed model may be an appropriate means for an insurance company to attract a low-risk customer segment.

Accident risks may further be reduced among adopters of the proposed insurance model because respondents, who intend to adopt it, stated a strong willingness to change their driving style. This willingness is even higher for those drivers who consider their current driving behavior as risky, and who perceive themselves as particularly cost sensitive. Overall, about half of our respondents indicated that they would be willing to change their driving style under the proposed insurance model. Consequently, the proposed insurance model may not only attract drivers who are already driving carefully, but can also motivate high-risk drivers to change their driving behavior. This socially desirable effect might be exploited to create positive public awareness for behavior-based automobile insurance contracts.

Our analysis of gender differences revealed that men and women follow different decision patterns with regard to their *Behavioral Intention to Adopt*. Although both genders show equal behavioral intentions, different factors determine their decision making. Women seem to have recognized that they may expect a larger monetary benefit due to their more defensive driving style and at the same time will experience fewer consequences for their driving pleasure. Men are less influenced by expected benefits but rather by their general attitude towards the proposed insurance product. From a marketing perspective, different strategies seem to be required for women and men. Women should be made aware of their potential to become rewarded for their already prevailing careful driving style. Social desirability might further improve the acceptance among women as *Social Influence* turned out to be significant only for women but not for men. Men might rather be motivated to adopt a behavior-based insurance scheme if they perceive it as generally positive and desirable. With regard to their *Relative Advantage*, awareness should be created that they get the possibility to influence their insurance premiums by adapting their driving style. Such a feeling of control might mitigate the relative disadvantage men initially perceive.

To summarize, the present study indicated that there might be a remarkable market segment for behavior-based automobile insurances that is so far underestimated. Several success factors could be derived from our analysis. Most important - and at the same time most critical - is the high level of expected monetary reward for careful driving. Gender differences have been revealed with regard to decision patterns of men and women, which should be incorporated in marketing strategies and the development of a product portfolio. From a societal perspective, our results indicate that behavior-based insurance models may be desirable due to their potential effect to motivate people to reduce their accident risks by changing their driving style in a positive way.

IV.7 Limitations

Although the presented empirical results indicate strong statistical precision and consistency and therefore corroborate the robustness of our findings, several critical points shall be discussed, which point the direction for further investigations in this research domain.

From a theoretical perspective, the objective of this study was to develop and empirically test a research model that is aimed at evaluating the perception and acceptance of persuasive business models. The model was tested with the example of behavior-based automobile insurance. Due to the fact that the proposed insurance model is neither a purely technical nor a pure product or service innovation, research models usually applied in information systems or marketing research could not be used directly. Therefore, constructs and measurement scales had to be adapted to the specific context of behavior-based automobile insurance. Although statistical reliability and validity criteria suggest a high level of precision for the measurement model, further empirical studies are required to analyze the external validity of the research model, i.e. its transferability and applicability to other contexts. Such contexts may be found primarily in business domains in which information asymmetries can be exploited by opportunistic behavior. Examples are other insurances, vehicle leasing, or the rental car business. Investigations in other business domains have to reveal whether our research model is only applicable to automobile insurance, or whether it is a valid instrument to analyze and explain the perception of the much broader concept of persuasive business models.

Furthermore, our sample was restricted to the German population. Cultural differences may lead to different results if our research model is applied in other regions. Therefore our results should be validated in other countries, and particularly in regions with a different cultural background. Potentially, such analyses will reveal different perception patterns, for example, between rather individualistic societies like in most European countries or the United States and rather collectivistic societies, which are prevailing in some Asian regions (Kitayama et al. 1997).

Whereas most constructs in our research model were based on literature from IS research, *Behavioral Intention to Change* represents a novel model extension that was added on the basis of theoretical considerations. The strong relationship with *Behavioral Intention to Adopt* has shown that in principle the assumed relationship could be confirmed. However, the moderate degree to which the variance of the BIC construct is explained raises the question which other factors may influence *Behavioral Intention to Change*. In the context of the investigation of persuasive business models, this question is relevant as it will help to predict which product or service characteristics may increase the effectiveness of persuasive business models with regard to their capability to achieve a desired behavioral change.

Due to the complexity of the proposed insurance model, a realistic field test could not be implemented. Instead, a textual scenario description was provided to explain the proposed concept to our respondents. Although this approach is widely applied in research practice, the predictive relevance of the stated *Behavioral Intention to Adopt* for actual behavior may be limited. Therefore field studies under realistic conditions should be conducted to analyze the degree to which the behavioral intentions investigated in our study are robust antecedents of actual behavior.

With regard to a potential commercialization of the proposed insurance model, our study has shown that consumers expect a relatively high saving potential. So far, it is not known whether decreasing accident risks or attracting low-risk customers would enable a monetary reward for careful drivers that is large enough to become broadly accepted. Field studies under realistic conditions are required to investigate to which degree behavior-based insurance models will motivate a more careful driving style and decrease accident risk. Based on these results, it can be estimated whether the expected premium savings can be granted to careful drivers.

V Conclusions

The starting point of this dissertation was the observation that persuasive technologies, which are intended to influence human behavior, become extended along three dimensions. First, formerly isolated technical artifacts are increasingly integrated into everyday technologies in the domestic domain to create so-called *persuasive environments*. Second, *profiling* as a means to adapt persuasive strategies to personality traits has recently become a fruitful research domain and may achieve commercial maturity in the near future. Third, *persuasive business models* have been developed to offer monetary rewards for desirable behavior. Although persuasive technologies have attracted strong focus in IS research, consumer perception and acceptance of persuasive environments and persuasive business models have so far been under-researched. Furthermore, a systematic evaluation of the effectiveness of profiling as an enabler for adaptive persuasive systems has not been conducted so far. The present dissertation fills these research gaps with three distinct empirical studies, the results of which are summarized in the following by answering the research questions raised at the beginning of this thesis.

> ***Q1:*** *Which factors influence the acceptance of a persuasive kitchen environment?*

Overall, the five scenarios that constitute the proposed persuasive kitchen environment achieved a moderately positive acceptance level. The strongest factor to influence the intention to adopt the proposed environment was *Performance Expectancy*, which indicates that consumers base their decision primarily on functional considerations. In contrast, *Effort Expectancy* plays only a minor role, which shows that our respondents had only minor concerns about the usability of the proposed scenarios. A surprisingly strong effect has been found for *Social Influence*, which may be attributable to the fact that the decision for or against the proposed kitchen environment is usually made in coordination with family members or cohabitants. In contrast to the majority of previous technology acceptance studies, gender actually moderates the influence of *Social Influence* such that men are more influenced by their social environment than women, whereas age has not shown any effect. *Importance* - the degree to which a person feels a certain need to get the proposed assistance - increases the influence of *Performance Expectancy* and decreases the influence of *Social Influence*. *Personal Relevance*, which measures a person's dedication and interest in the analyzed domain, has turned out to have no effect. Finally, *Personal Innovativeness in IT* has been shown to increase the effect of *Social Influence*.

Furthermore, our analysis has shown that persuasiveness is not regarded as the key feature of the proposed kitchen environment. Convenience and joyfulness seem to be stronger arguments for an adoption decision than the expectation of being supported in adopting more healthy nutrition habits. This makes us conclude that a buying decision for a persuasive kitchen environment would not be based merely on the desire to lead a healthier lifestyle. At least equally important is the expectation that cooking becomes more convenient and joyful.

> ***Q2:*** *Can profiling improve the effectiveness of persuasive messages?*

The effectiveness of persuasive technologies largely varies with the mode of how persuasive principles are selected. In the case where the persuasive principle is incongruent with the personality traits of a subject, no persuasive effect could be achieved, whereas applying the best-fitting principle induces the desired change in behavior. However, no significant difference could be observed between a random selection of persuasive principles and the application of the most appropriate principle. Consequently, profiling can improve the effectiveness of persuasive messages compared to the proliferation of single-strategy messages, but it is not superior to a random selection of messages that adopt different persuasive strategies. Although further investigations are necessary to corroborate these findings, our results indicate that the technical effort for implementing an adaptive persuasive system may be mitigated without a loss of effectiveness by randomly mixing different persuasive principles.

> ***Q3:*** *Which factors influence the acceptance of behavior-based automobile insurance and their effect on the willingness to change driving behavior?*

A relatively high level of acceptance could be observed for the proposed behavior-based automobile insurance. The expected monetary benefit turned out to be the major factor to influence the intention to adopt this novel insurance model. Other important factors are confidence in the technology (*Performance Expectancy)*, the expected influence on driving pleasure (*Perceived Enjoyment)*, and the general *Attitude* towards the insurance model. *Trust* in the insurance company and *Perceived Privacy* issues play only a minor yet significant role. *Social Influence* has no direct effect but a strong indirect effect on acceptance, which is mediated by *Attitude.* Although behavior-based insurance models increase the complexity of automobile insurance, *Effort Expectancy* has no significant influence on an adoption decision. *Gender* has a moderating effect on several relationships in the research model. The effects of *Social Influence, Perceived Enjoyment,* and *Relative Advantage* are stronger for women than for men, whereas the effect of *Attitude* is stronger for men than for women. As the expected monetary benefit is the strongest influencing factor, it is of particular interest to know how much financial compensation consumers expect in order to accept a behavior-based insurance model. In presence of a reference premium of 700 EUR per year for a conventional contract, a potential premium reduction of at least 200 EUR greatly increases the acceptance of the proposed insurance model compared to lower compensation levels.

With regard to the effect of behavior-based automobile insurance on the willingness to change driving behavior, a majority of respondents state that they would change their driving behavior under a behavior-based insurance contract. Three factors influence the willingness to change driving behavior under the proposed insurance regime, namely the general level of acceptance, cost sensitivity with regard to the total costs of vehicle ownership, and driving style. People, who have stated higher acceptance levels, or are more cost sensitive, or have a more disadvantageous driving style, show a stronger willingness to change their driving behavior. Consequently, our results indicate that behavior-based insurance models can be expected to achieve a change of driving behavior among their adopters.

In summary, the three studies have shown that persuasive environments and persuasive business models can be expected to gain substantial market acceptance, and both could be a promising way for a more widespread proliferation of persuasive technologies. Furthermore, attention should be paid to an appropriate selection and implementation of persuasive strategies, since either applying a broad set of different strategies or adapting persuasive principles to individual personality traits can be expected to enhance the effectiveness of a persuasive system.

References

Agarwal, R., and J. Prasad (1998), "A Conceptual and Operational Definition of Personal Innovativeness in the Domain of Information Technology," *Information Systems Research*, 9(2), 204–15.

Ahuja, M.K., and J.B. Thatcher (2005), "Moving Beyond Intentions and Toward the Theory of Trying: Effects of Work Environment and Gender on Post-Adoption Information Technology Use," *MIS Quarterly*, 29(3), 427–59.

Aiken, L.S., and S.G. West (1991), *Multiple Regression: Testing and Interpreting Interactions*, London: Sage.

Ainuddin, R., P.W. Beamish, J.S. Hulland, and M.J. Rouse (2007), "Resource Attributes and Firm Performance in International Joint Ventures," *Journal of World Business*, 42(1), 47–60.

Ajzen, I. (1991), "The Theory of Planned Behavior," *Organizational Behavior and Human Decision Processes*, 50(2), 179–211.

Akerlof, G.A. (1970), "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism," *The Quarterly Journal of Economics*, 84(3), 488–500.

Alin, A. (2010), "Multicollinearity," *Wiley Interdisciplinary Reviews: Computational Statistics*, 2(3), 370–74.

Arrow, K.J. (1963), "Uncertainty and the Welfare Economics of Medical Care," *The American Economic Review*, 53(5), 941–73.

Arteaga, S.M., M. Kudeki, and A. Woodworth (2009), "Combating Obesity Trends in Teenagers through Persuasive Mobile Technology," *ACM SIGACCESS Accessibility and Computing*, 2(94), 17–25.

Atkinson, M.A., and C. Kydd (1997), "Individual Characteristics Associated with World Wide Web Use: an Empirical Study of Playfulness and Motivation," *ACM SIGMIS Advances in Information Systems*, ACM, 28(2), 53–62.

Backhaus, K., B. Blechschmidt, and M. Eisenbeiss (2006), "Der Stichprobeneinfluss bei Kausalanalysen," *Die Betriebswirtschaft*, 66(6), 711–27.

Backhaus, K., B. Erichson, W. Plinke, and R. Weiber (2011), *Multivariate Analysemethoden - Eine anwendungsorientierte Einführung*, Berlin / Heidelberg: Springer.

Bagozzi, R.P. (1979), "The Role of Measurement in Theory Construction and Hypothesis Testing: Towards a Holistic Model," in *Conceptual and Theoretical Developments in Marketing*, O. Ferrel, S. Brown, and C. Lamb, eds., Chicago: American Marketing Association, 15–32.

——— (1980), *Causal Models in Marketing*, New York: Wiley.

Bagozzi, R.P., and L.W. Phillips (1982), "Representing and Testing Organizational Theories: A Holistic Construal," *Administrative Science Quarterly*, 27(3), 459–89.

Barki, H., and J. Hartwick (1994), "Measuring User Participation, User Involvement, and User Attitude," *MIS Quarterly*, 18(1), 59–82.

Bauer, J., K. Streefkerk, and R.R. Varick (2005), "Fridgets: Digital Refrigerator Magnets," in *Proceedings of the Conference on Human Factors in Computing Systems*, Portland, USA: ACM, 2060–64.

Bem, S.L. (1981), "Gender Schema Theory: A Cognitive Account of Sex Typing," *Psychological Review*, 88(4), 354–64.

Bernardin, H.J. (1977), "Behavioral Expectation Scales versus Summated Scales: A Fairer Comparison," *Journal of Applied Psychology*, 62(4), 422–27.

Bliemel, F., A. Eggert, G. Fassott, and J. Henseler (2005), "PLS und Kovarianzstrukturanalyse im Vergleich," in *Handbuch PLS-Pfadmodellierung. Methode, Anwendung, Praxisbeispiele*, F. Bliemel, A. Eggert, G. Fassott, and J. Henseler, eds., Stuttgart: Schäffer-Poeschel, 9–16.

Blumendorf, M., S. Feuerstack, and S. Albayrak (2008), "Multimodal User Interfaces for Smart Environments: the Multi-access Service Platform," in *Proceedings of the Working Conference on Advanced Visual Interfaces*, Napoli, Italy, 478–79.

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

Boomsma, A. (1982), "The Robustness of LISREL against Small Sample Size in Factor Analysis Models," in *Systems Under Direct Observations: Causality, Structure, Prediction - Part 1*, K. G. Jöreskog and H. Wold, eds., Amsterdam: North-Holland, 149–74.

Bordoff, J., and P.J. Noel (2008), "Pay-As-You-Drive Auto Insurance: A Simple Way to Reduce Driving-Related Harms and Increase Equity," The Brookings Institution.

Bortz, J., and C. Schuster (2010), *Statistik für Human- und Sozialwissenschaftler*, Berlin: Springer.

Brown, S.A., and V. Venkatesh (2005), "Model of Adoption of Technology in Households: A Baseline Model Test and Extension Incorporating Household Life Cycle," *MIS Quarterly*, 29(3), 399–426.

Brown, T.A. (2006), *Confirmatory Factor Analysis for Applied Research*, New York: The Guilford Press.

Cacioppo, J.T., R.E. Petty, C.F. Kao, and R. Rodriguez (1986), "Central and Peripheral Routes to Persuasion: An Individual Difference Perspective," *Journal of Personality and Social Psychology*, 51(5), 1032–43.

Carli, L.L. (2001), "Gender and Social Influence," *Journal of Social Issues*, 57(4), 725–41.

Carlson, R. (1971), "Sex Differences in Ego Functioning," *Journal of Consulting and Clinical Psychology*, 37(2), 267–77.

Carlsson, C., J. Puhakainen, J. Carlsson, P. Walden, and K. Hyvonen (2006), "Adoption of Mobile Devices/Services - Searching for Answers with the UTAUT," in *Proceedings of the 39th Annual Hawaii International Conference on System Sciences*, Washington: IEEE Computer Society.

Carte, T.A., and C.J. Russell (2003), "In Pursuit of Moderation: Nine Common Errors and their Solutions," *MIS Quarterly*, 27(3), 479–501.

Cassel, C., P. Hackl, and A.H. Westlund (1999), "Robustness of Partial Least-Squares Method for Estimating Latent Variable Quality Structures," *Journal of Applied Statistics*, 26(4), 435–46.

Chellapa, R.K., and R.G. Sin (2005), "Personalization versus Privacy: An Empirical Examination of the Online Consumer's Dilemma," *Information Technology and Management*, 6(2-3), 181–202.

Chin, W.W. (1998a), "The Partial Least Squares Approach for Structural Equation Modeling," in *Modern Methods for Business Research*, G. A. Marcoulides, ed., Mahwah, USA: Lawrence Erlbaum Associates, 295–336.

——— (1998b), "Commentary: Issues and Opinion on Structural Equation Modeling," *MIS Quarterly*, 22(1), vii-xvi.

Chin, W.W., B.L. Marcolin, and P.R. Newsted (2003), "A Partial Least Squares Latent Variable Modeling Approach for Measuring Interaction Effects: Results from a Monte Carlo Simulation Study and an Electronic-Mail Emotion/Adoption Study," *Information Systems Research*, 14(2), 189–217.

Chin, W.W., and J. Dibbern (2010), "An Introduction to a Permutation Based Procedure for Multi-group PLS Analysis: Results of Tests of Differences on Simulated Data and a Cross Cultural Analysis of the Sourcing of Information System Services between Germany and the USA," in *Handbook of Partial Least Squares*, V. E. Vinzi, W. W. Chin, J. Henseler, and H. Wang, eds., Berlin / Heidelberg: Springer, 171–93.

Chin, W.W., and P.R. Newsted (1999), "Structural Equation Modeling Analysis with Small Samples Using Partial Least Squares," in *Statistical Strategies for Small Sample Research*, R. Hoyle, ed., Thousand Oaks, USA, 307–41.

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

——— (1988), *Basic Marketing Research*, Chicago: The Dryden Press.

Cialdini, R.B. (2001), "Harnessing the Science of Persuasion," *Harvard Business Review*, 79(9), 72–79.

——— (2008), *Influence, Science and Practice*, Boston: Prentice Hall.

Cialdini, R.B., L.J. Demaine, B.J. Sagarin, D.W. Barrett, and P.L. Winter (2006), "Managing Social Norms for Persuasive Impact," *Social Influence*, 1(1), 3–15.

Cialdini, R.B., C.A. Kallgren, and R.R. Reno (1991), "A Focus Theory of Normative Conduct: A Theoretical Refinement and Reevaluation of the Role of Norms in Human Behavior," *Advances in Experimental Social Psychology*, 24, 201–34.

Cialdini, R.B., M.R. Trost, and J.T. Newsom (1995), "Preference for Consistency: The Development of a Valid Measure and the Discovery of Surprising Behavioral Implications," *Journal of Personality and Social Psychology*, 69(2), 318–28.

Cialdini, R.B., W. Wosinska, D.W. Barrett, J. Butner, and M. Gornik-Durose (1999), "Compliance with a Request in Two Cultures: The Differential Influence of Social Proof and Commitment/Consistency on Collectivists and Individualists," *Personality and Social Psychology Bulletin*, 25(10), 1242–53.

Cialdini, R.B., and N.J. Goldstein (2004), "Social Influence: Compliance and Conformity," *Annual Review of Phsychology*, 55, 591–621.

Cohen, J. (1988), *Statistical Power Analysis for the Behavioral Sciences*, Hillsdale, USA: Lawrence Erlbaum Associates.

Consolvo, S., P. Klasnja, D.W. McDonald, and J.A. Landay (2009), "Goal-setting Considerations for Persuasive Technologies that Encourage Physical Activity," in *Proceedings of the 4th International Conference on Persuasive Technology*, New York: ACM Press.

Conway, J.M., and A.I. Huffcutt (2003), "A Review and Evaluation of Exploratory Factor Analysis Practices in Organizational Research," *Organizational Research Methods*, 6(2), 147–68.

Cook, M., and D. T. Campbell (1979), *Quasi-Experimentation: Design & Analysis Issues for Field Settings*, Boston, USA: Houghton Mifflin.

Coxon, A.P.M. (1999), *Sorting Data. Collection and Analysis*, Thousand Oaks: Sage.

Crawford, M., R. Chaffin, and L. Fitton (1995), "Cognition in Social Context," *Learning and Individual Differences*, 7(4), 341–62.

Dabholkar, P.A., and R.P. Bagozzi (2002), "An Attitudinal Model of Technology-based Self-Service: Moderating Effects of Consumer Traits and Situational Factors," *Journal of the Academy of Marketing Science*, 30(3), 184–201.

Dahlby, B.G. (1983), "Adverse Selection and Statistical Discrimination: An Analysis of Canadian Automobile Insurance," *Journal of Public Economics*, 20(1), 121–30.

Davis, F.D. (1989), "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology," *MIS Quarterly*, 13(3), 319–40.

Davis, F.D., R.P. Bagozzi, and P.R. Warshaw (1992), "Extrinsic and Intrinsic Motivation to Use Computers in the Workplace," *Journal of Applied Psychology*, 22(14), 1111–32.

Diamantopoulos, A., P. Riefler, and K.P. Roth (2008), "Advancing Formative Measurement Models," *Journal of Business Research*, 61(12), 1203–18.

Diamantopoulos, A., and M. Winklhofer (2001), "Index Construction with Formative Indicators: An Alternative to Scale Development," *Journal of Marketing*, 38(2), 269–77.

Dillman, D.A. (2007), *Mail and Internet Surveys: The Tailored Design Method*, Hoboken, USA: John Wiley & Sons.

Dinev, T., and P. Hart (2006), "An Extended Privacy Calculus Model for E-Commerce Transactions," *Information Systems Research*, 17(1), 61–80.

D'Arcy, S.P., and N.A. Doherty (1990), "Adverse Selection, Private Information, and Lowballing in Insurance Markets," *The Journal of Business*, 63(2), 145–64.

Eagly, A.H. (1978), "Sex Differences in Influenceability," *Psychological Bulletin*, 85(1), 86–116.

Eastin, M. (2002), "Diffusion of e-commerce: an Analysis of the Adoption of Four e-commerce Activities," *Telematics and Informatics*, 19(3), 251–67.

Edlin, A.S. (2003), "Per-Mile Premiums for Auto Insurance," in *Economics for an Imperfect World: Essays in Honor of Joseph Stiglitz*, R. Arnott, B. Greenwald, and R. Kanbur, eds., Cambridge, USA: MIT Press, 55–82.

Faul, F., E. Erdfelder, A. Buchner, and A.G. Lang (2009), "Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses," *Behavior Research Methods*, 41(4), 1149–60.

Ferneley, E., and B. Light (2008), "Unpacking User Relations in an Emerging Ubiquitous Computing Environment: Introducing the Bystander," *Journal of Information Technology*, 23(3), 163–75.

Field, A. (2009), *Discovering Statistics Using SPSS*, London: Sage.

Filipova-Neumann, L., and P. Welzel (2010), "Reducing Asymmetric Information in Insurance Markets: Cars with Black Boxes," *Telematics and Informatics*, 27(4), 394–403.

Fishbein, M., and I. Ajzen (1975), *Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research*, Reading, USA: Addison-Wesley.

Fitzmaurice, G.M., N.M. Laird, and J.H. Ware (2004), *Applied Longitudinal Analysis*, Hoboken, USA: Wiley.

Fogg, B.J. (1999), "Persuasive Technologies," *Communications of the ACM*, 42(5), 26–29.

——— (2002), *Persuasive Technology*, San Francisco: Morgan Kaufmann.

Fogg, B.J., and J. Hreha (2010), "Behavior Wizard: A Method for Matching Target Behaviors with Solutions," in *Proceedings of Persuasive 2010*, T. Ploug, P. Hasle, and H. Oinas-Kukkonen, eds., Berlin / Heidelberg: Springer, 117–31.

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

Froehlich, J., T. Dillahunt, P. Klasnja, J. Mankoff, S. Consolvo, B. Harrison, and J.A. Landay (2009), "UbiGreen: Investigating a Mobile Tool for Tracking and Supporting Green Transportation Habits," in *Proceedings of the 27th International Conference on Human Factors in Computing Systems*, Boston: ACM, 1043–52.

Garai, J., and A. Scheinfeld (1968), "Sex Differences in Mental and Behavioral Traits," *Genetic Psychology Monographs*, 77(2), 169–299.

Garfield, M.J. (2005), "Acceptance of Ubiquituous Computing," *Information Systems Management*, 22(4), 24–31.

Gefen, D., E. Karahanna, and D. Straub (2003), "Trust and TAM in Online Shopping: An Integrated Model," *Management Information Systems*, 27(1), 51–90.

Gefen, D., D.W. Straub, and M.C. Boudreau (2000), "Structural Equation Modeling and Regression: Guidelines for Research Practice," *Communications of the Association for Information Systems*, 4(7), 1–77.

Gefen, D., and D.W. Straub (2004), "Consumer Trust in B2C e-Commerce and the Importance of Social Presence: Experiments in e-Products and e-Services," *Omega*, 32(6), 407–24.

Geisser, S. (1974), “A Predictive Approach to the Random Effect Model,” *Biometrika*, 61(1), 101–7.

Goldberg, L.R. (1993), “The Structure of Phenotypic Personality Traits,” *The American Psychologist*, 48(1), 26–34.

Goldstein, N.J., R.B. Cialdini, and V. Griskevicius (2008), “A Room with a Viewpoint: Using Social Norms to Motivate Environmental Conservation in Hotels,” *Journal of Consumer Research*, 35(3), 472–82.

Grewal, R., J.A. Cote, and H. Baumgartner (2004), “Multicollinearity and Measurement Error in Structural Equation Models: Implications for Theory Testing,” *Marketing Science*, 23(4), 519–29.

Grice, J.W. (2001), “Computing and Evaluating Factor Scores,” *Psychological Methods*, 6(4), 430–50.

Grice, J.W., and R.J. Harris (1998), “A Comparison of Regression and Loading Weights for the Computation of Factor Scores,” *Multivariate Behavioral Research*, 33(2), 221–47.

Grolleman, J., B. van Dijk, A. Nijholt, and A. van Emst (2006), “Break the Habit! Designing an e-Therapy Intervention Using a Virtual Coach in Aid of Smoking Cessation,” in *Persuasive Technology*, W. IJsselsteijn, Y. de Kort, Cees Midden, B. Eggen, and E. van den Hoven, eds., Berlin / Heidelberg: Springer, 133–41.

Guadagno, R.E., T. Asher, L.J. Demaine, and R.B. Cialdini (2001), “When Saying Yes Leads to Saying No: Preference for Consistency and the Reverse Foot-in-the-Door Effect,” *Personality and Social Psychology Bulletin*, 27(7), 859–67.

Guensler, R., A. Amekudzi, and J. Williams (2003), “Current State Regulatory Support for Pay-As-You-Drive Automobile Insurance Options,” *Journal of Insurance Regulation*, 21(3), 31–51.

Gueorguieva, R., and J.H. Krystal (2004), “Move over ANOVA: Progress in Analyzing Repeated-Measures Data and its Reflection in Papers Published in the Archives of General Psychiatry,” *Archives of General Psychiatry*, 61(3), 310–17.

Gujarati, D.N. (2003), *Basic Econometrics*, Boston: McGraw-Hill.

Gurka, M.J. (2006), “Selecting the Best Linear Mixed Model under REML,” *The American Statistician*, 60(1), 19–26.

Haines, M., and S.F. Spear (1996), “Changing the Perception of the Norm: A Strategy to Decrease Binge Drinking among College Students,” *Journal of American College Health*, 45(3), 134–40.

Hair, J.F., W.C. Black, B.J. Babin, R.E. Anderson, and R.L. Tatham (2006), *Multivariate Data Analysis*, Upper Saddle River, USA: Prentice-Hall.

Hair, J.F., C.M. Ringle, and M. Sarstedt (2011), "PLS-SEM: Indeed a Silver Bullet," *The Journal of Marketing Theory and Practice*, 19(2), 139–52.

Haitovsky, Y. (1969), "Multicollinearity in Regression Analysis: A Comment," *Review of Economics and Statistics*, 51(4), 486–89.

Halko, S., and J. Kientz (2010), "Personality and Persuasive Technology: An Exploratory Study on Health-Promoting Mobile Applications," in *Proceedings of the 5th International Conference on Persuasive Technology*, T. Ploug, P. Hasle, and H. Oinas-Kukkonen, eds., Berlin / Heidelberg: Springer, 150–61.

Hansen, T. (2005), "Perspectives on Consumer Decision Making: An Integrated Approach," *Journal of Consumer Behaviour*, 4(6), 420–37.

Hanson-Smith, V., D. Wimalasuriya, and A. Fortier (2006), "NutriStat: Tracking Young Child Nutrition," in *Proceedings of the Conference on Human Factors in Computing Systems*, Montreal, Canada: ACM, 1831–36.

Harloff, J. (2005), "Multiple Level Weighted Card Sorting," *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences*, 1(4), 119–28.

Harré, N., J. Field, and B. Kirkwood (1996), "Gender Differences and Areas of Common Concern in the Driving Behaviors and Attitudes of Adolescents," *Journal of Safety Research*, 27(3), 163–73.

Heijden, H. van der (2004), "User Acceptance of Hedonic Information Systems," *MIS Quarterly*, 28(4), 695–704.

Hennessy, D.A., and D.L. Wiesenthal (2001), "Gender, Driver Aggression, and Driver Violence: An Applied Evaluation," *Sex Roles*, 44(11), 661–76.

Henseler, J. (2007), "A New and Simple Approach to Multi-Group Analysis in Partial Least Squares Path Modeling," in *Proceedings of PLS'07: 5th International Symposium on PLS and Related Methods*, Matforsk, Norway: As, 104–7.

Henseler, J., C.M. Ringle, and R.R. Sinkovics (2009), "The Use of Partial Least Squares Modeling in International Marketing," *Advances in International Marketing*, 20, 277–319.

Henseler, J., and G. Fassott (2010), "Testing Moderating Effects in PLS Path Models: An Illustration of Available Procedures," in *Handbook of Partial Least Squares*, V. E. Vinzi, W. W. Chin, J. Henseler, and H. Wang, eds., Berlin / Heidelberg: Springer, 713–35.

Herrmann, A., F. Huber, and F. Kressmann (2006), “Varianz- und kovarianzbasierte Strukturgleichungsmodelle - Ein Leitfaden zu deren Spezifikation, Schätzung und Beurteilung,” *Zeitschrift für betriebswirtschaftliche Forschung*, 58(2), 34–66.

Hoffman, L.W. (1972), “Early Childhood Experiences and Women’s Achievement Motives,” *Journal of Social Issues*, 28(2), 129–55.

Holt, D.B. (1995), “How Consumers Consume: A Typology of Consumption Practices,” *Journal of Consumer Research*, 22(1), 1–16.

Holzmüller, H.H., and H. Kasper (1991), “On a Theory of Export Performance: Personal and Organizational Determinants of Export Trade Activities Observed in Small and Medium-sized Firms,” *Management International Review*, 31(Special Issue), 45–70.

Homburg, C., and A. Giering (1996), “Konzeptualisierung und Operationalisierung komplexer Konstrukte. Ein Leitfaden für die Marketingforschung,” *Marketing ZFP*, 1(1), 5–24.

Homburg, C., and M. Klarmann (2006), “Die Kausalanalyse in der empirischen betriebswirtschaftlichen Forschung - Problemfelder und Anwendungsempfehlungen,” *Die Betriebswirtschaft*, 66(6), 727–48.

Hong, S.J., J.Y.L. Thong, J.Y. Moon, and K.Y. Tam (2008), “Understanding the Behavior of Mobile Data Services Consumers,” *Information Systems Frontiers*, 10(4), 431–45.

Hong, W., J.Y.L. Thong, W.M. Wong, and K.Y. Tam (2002), “Determinants of User Acceptance of Digital Libraries: an Empirical Examination of Individual Differences and System Characteristics,” *Journal of Management Information Systems*, 18(3), 97–124.

Hu, L.T., and P.M. Bentler (1998), “Fit Indices in Covariance Structure Modeling: Sensitivity to Underparameterized Model Misspecification,” *Psychological Methods*, 3(4), 424–53.

Hulland, J. (1999), “Use of Partial Least Squares (PLS) in Strategic Research: a Review of Four Recent Studies,” *Strategic Management Journal*, 20(2), 195–204.

Hutcheson, G., and N. Sofroniou (1999), *The Multivariate Social Scientist*, London: Sage.

Igbaria, M., S. Parasuraman, and J.J. Baroudi (1996), “A Motivational Model of Microcomputer Usage,” *Journal of Management Information Systems*, 13(1), 127–43.

Igbaria, M., S.J. Schiffman, and T.J. Wieckowski (1994), "The Respective Roles of Perceived Usefulness and Perceived Fun in the Acceptance of Microcomputer Technology," *Behaviour & Information Technology*, 13(6), 349–61.

Ippisch, T. (2010), *Telematics Data in Motor Insurance: Creating Value by Understanding the Impact of Accidents on Vehicle Use*, USA: Lulu Enterprises.

Jaccard, J., R. Turrisi, and C.K. Wan (1990), *Interaction Effects in Multiple Regression*, *Quantitative Applications in the Social Sciences (72)*, Newbury Park, USA: Sage.

Jarvenpaa, S.L., and P.A. Toad (1996), "Consumer Reactions to Electronic Shopping on the World Wide Web," *International Journal of Electronic Commerce*, 1(2), 59–88.

Johnson, M.D., A. Herrmann, and F. Huber (2006), "The Evolution of Loyalty Intentions," *Journal of Marketing*, 70(2), 122–32.

Ju, W., R. Hurwitz, T. Judd, and B. Lee (2001), "CounterActive: An Interactive Cookbook for the Kitchen Counter," in *Proceedings of the Conference on Human Factors in Computing Systems*, New York: ACM New York, 269–70.

Jöreskog, K.G., and D. Sörbom (1982), "Recent Developments in Structural Equation Modeling," *Journal of Marketing Research*, 19(4), 404–16.

Jöreskog, K.G., and H.O. Wold (1982), "The ML and PLS Technique for Modelling with Latent Variables: Historical and Comparative Aspects," in *Systems under Indirect Observation, Part I*, K. G. Jöreskog and H. O. Wold, eds., Amsterdam: North-Holland, 263–70.

Kaptein, M., J. Lacroix, and P. Saini (2010), "Individual Differences in Persuadability in the Health Promotion Domain," in *Proceedings of the 5th International Conference on Persuasive Technology*, T. Ploug, P. Hasle, and H. Oinas-Kukkonen, eds., Springer Berlin / Heidelberg, 94–105.

Kaptein, M., P. Markopoulos, B. de Ruyter, and E. Aarts (2009), "Can You Be Persuaded? Individual Differences in Susceptibility to Persuasion," in *Proceedings of Interact 2009*, T. Gross, J. Gulliksen, P. Kotzé, L. Oestreicher, P. Palanque, R. Prates, and M. Winckler, eds., Springer Berlin / Heidelberg, 115–18.

Kaptein, M., and D. Eckles (2010a), "Selecting Effective Means to Any End: Futures and Ethics of Persuasion Profiling," in *Persuasive Technology*, T. Ploug, P. Hasle, and H. Oinas-Kukkonen, eds., Berlin / Heidelberg: Springer, 82–93.

——— (2010b), "Persuasion Profiling: Identifying and Adapting to Individual Differences in Influence Strategy Effectiveness," Unpublished Work.

Keil, M., B. Tan, K.K. Wei, T. Saarinen, V. Tuunainen, and A. Wassenaar (2000), "A Cross-Cultural Study on Escalation of Commitment Behavior in Software Projects," *MIS Quarterly*, 24(2), 299–325.

Kellermann, K., and T. Cole (1994), "Classifying Compliance Gaining Messages: Taxonomic Disorder and Strategic Confusion," *Communication Theory*, 4(1), 3–60.

Kelman, H.C., and C.I. Hovland (1953), "'Reinstatement' of the Communicator in Delayed Measurement of Opinion Change," *Journal of Abnormal and Social Psychology*, 48(3), 327–35.

Kidd, C.D., R.J. Orr, G.D. Abowd, C.G. Atkeson, I.A. Essa, and B. MacIntyre (1999), "The Aware Home: A Living Laboratory for Ubiquitous Computing Research," in *Cooperative Buildings: Integrating Information, Organizations, and Architecture*, N. . Streita, J. Siegel, V. Harkopf, and K. S., eds., Berlin: Springer, 191–98.

King, W.R., and J. He (2006), "A Meta-Analysis of the Technology Acceptance Model," *Information & Management*, 43(6), 740–55.

Kinicki, A.J., J.E. Heyl, and T.E. Callarman (1986), "Assessing the Validity of the Cox, Zmud, and Clark Material-Requirements-Planning Audit Instrument," *The Academy of Management Journal*, 29(3), 633–41.

Kitayama, S., H.R. Markus, R. Matsumoto, and V. Norasakkunit (1997), "Individual and Collective Processes in the Construction of the Self: Self-Enhancement in the United Stated and Self-Criticism in Japan," *Journal of Personality and Social Psychology*, 72, 1245–67.

Kline, P. (1999), *The Handbook of Psychological Testing*, London: Routledge.

Koivumäki, T., A. Ristola, and M. Kesti (2006), "The Perceptions towards Mobile Services: an Empirical Analysis of the Role of Use Facilitators," *Personal and Ubiquitous Computing*, 12(1), 67–75.

Krafft, M., O. Götz, and K. Liehr-Gobbers (2005), "Die Validierung von Strukturgleichungsmodellen mit Hilfe des Partial-Least-Squares (PLS)-Ansatz," in *Handbuch PLS-Pfadmodellierung. Methode, Anwendung, Praxisbeispiele*, F. Bliemel, A. Eggert, G. Fassott, and G. Henseler, eds., Stuttgart: Schäffer-Poeschel, 71–116.

Kreuter, M.W., D. Farrell, L. Olevitch, and L. Brennan (2000), *Tailoring Health Messages: Customizing Communication with Computer Technology*, Mahwah, USA: Erlbaum.

Lacroix, J., P. Saini, and A. Goris (2009), "Understanding User Cognitions to Guide the Tailoring of Persuasive Technology-based Physical Activity Interventions," in *Proceedings of the 4th International Conference on Persuasive Technology*, New York: ACM, 9:1–9:8.

Landwehr, J.R., A. Herrmann, and M. Heitmann (2008), "Linear Mixed Models. Grundidee, Methodik und Anwendung," *Marketing ZFP*, 30(3), 175–86.

Lee, H., W. Lee, and Y.K. Lim (2010), "The Effect of Eco-driving System towards Sustainable Driving Behavior," in *Proceedings of the 28th International Conference on Human Factors in Computing Systems*, New York: ACM, 4255–60.

Lee, Y., K.A. Kozar, and K.R.T. Larsen (2003), "The Technology Acceptance Model: Past, Present, and Future," *Communications of the Association for Information Systems*, 12(50), 752–80.

Levine, T., and R. Wheeless (1990), "Cross-Situational Consistency and Use/Nonuse Tendencies in Compliance-gaining Tactic Selection," *Southern Communication Journal*, 56(1), 1–11.

Lin, J., L. Mamykina, S. Lindtner, G. Delajoux, and H. Strub (2006), "Fish ' n ' Steps: Encouraging Physical Activity with an Interactive Computer Game," in *Proceeding of UbiComp 2006*, P. Dourish and A. Friday, eds., Berlin / Heidelberg: Springer, 261–78.

Lindberg, G., L. Hultkrantz, J.E. Nilsson, and F. Thomas (2005), "Pay-As-You-Speed: Two Field Experiments on Controlling Adverse Selection and Moral Hazard in Traffic Insurance," Göteborgs Universitet.

Litman, T. (2008), "Distance-based Vehicle Insurance: Feasibility, Costs and Benefits - Comprehensive Technical Report," Victoria Transport Policy Institute (VTPI).

Lo, J.L., T.Y. Lin, H.H. Chu, H.C. Chou, and J.H. Chen (2007), "Playful Tray: Adopting Ubicomp and Persuasive Techniques into Play-based Occupational Therapy for Reducing Poor Eating Behavior in Young Children," in *Proceedings of UbiComp 2007*, J. Krumm, G. Abowd, A. Seneviratne, and T. Strang, eds., Berlin / Heidelberg: Springer.

Lohmöller, J.B. (1989), *Latent Variable Path Modeling with Partial Least Squares*, Heidelberg: Physica-Verlag.

Loock, C.-M., T. Staake, and J.R. Landwehr (2011), "Green IS Design and Energy Conservation: An Empirical Investigation of Social Normative Feedback," in *Proceedings of ICIS 2011*.

Luarn, P., and H. Lin (2005), "Toward an Understanding of the Behavioral Intention to Use Mobile Banking," *Computers in Human Behavior*, 21(6), 873–91.

Lundeberg, M.A., P.W. Fox, and J. Puncochar (1994), "Highly Confident but Wrong: Gender Differences and Similarities in Confidence Judgments," *Journal of Educational Psychology*, 86(1), 114–21.

Maccoby, E.E., and C.N. Jacklin (1974), *The Psychology of Sex Differences*, Stanford, USA: Stanford University Press.

Malhotra, N.K., and D.F. Birks (2003), *Marketing Research: An Applied Approach*, New Jersey, USA: Prentice Hall.

Mallat, N., M. Rossi, V.K. Tuunainen, and A. Öörni (2006), "An Empirical Investigation of Mobile Ticketing Service Adoption in Public Transportation," *Personal and Ubiquitous Computing*, 12(1), 57–65.

Marcoulides, G.A., W.W. Chin, and C. Saunders (2009), "A Critical Look at Partial Least Squares Modeling," *MIS Quarterly*, 33(1), 171–75.

Marsh, H.W., J.R. Balla, and K.-T. Hau (1996), "An Evaluation of Incremental Fit Indices: A Clarification of Mathematical and Empirical Properties," in *Advanced Structural Equation Modeling*, G. A. Marcoulides and R. E. Schuhmacker, eds., Mahwah, USA: Lawrence Erlbaum Associates, 315–53.

Marwell, G., and D.R. Schmitt (1967), "Dimensions of Compliance-Gaining Behavior: An Empirical Analysis," *Sociometry*, 30(4), 350–64.

Mathwick, C., N. Malhotra, and E. Rigdon (2001), "Experiential Value: Conceptualization , Measurement and Application in the Catalog and Internet Shopping Environment," *Journal of Retailing*, 77, 39–56.

McKnight, D.H., V. Choudhury, and C. Kacmar (2002), "Developing and Validating Trust Measures for e-Commerce: An Integrative Typology," *Information Systems Research*, 13(3), 334–59.

Meschtscherjakov, A., D. Wilfinger, T. Scherndl, and M. Tscheligi (2009), "Acceptance of Future Persuasive in-car Interfaces towards a more Economic Driving Behaviour," in *Proceedings of the 1st International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, New York: ACM, 81–88.

Meyers-Levy, J., and D. Maheswaran (1991), "Exploring Differences in Males' and Females' Processing Strategies," *Journal of Consumer Research: An Interdisciplinary Quarterly*, 18(1), 63–70.

Midden, C., and J. Ham (2009), "Using Negative and Positive Social Feedback from a Robotic Agent to Save Energy," in *Proceedings of the 4th International Conference on Persuasive Technology*, New York: ACM, 12:1--12:6.

Minton, C., J Kagan, and J.A. Levine (1971), "Maternal Control and Obedience in the Two-Year-Old," *Child Development*, 42(6), 1873–94.

Minton, H.L., and F.W. Schneider (1980), *Differential Psychology*, Prospect Heights, USA: Waveland Press.

Moon, J.-W., and Y.-G. Kim (2001), "Extending the TAM for a World-Wide-Web context," *Information & Management*, 38(4), 217–30.

Moon, Y. (2002), "Personalization and Personality: Some Effects of Customizing Message Style Based on Consumer Personality," *Journal of Consumer Psychology*, 12(4), 313–26.

Moore, G.C., and I. Benbasat (1991), "Development of an Instrument to Measure the Perceptions of Adopting an Information Technology Innovation," *Information Systems Research*, 2(3), 192–222.

Morris, M.G., and V. Venkatesh (2000), "Age Differences in Technology Adoption Decisions: Implications for a Changing Work Force," *Personnel Psychology*, 53(2), 375–403.

Mukherjee, A., and P. Nath (2007), "Role of Electronic Trust in Online Retailing: A Re-Examination of the Commitment-Trust Theory," *European Journal of Marketing*, 41(9/10), 1173–1202.

Noar, S.M., C.N. Benac, and M.S. Harris (2007), "Does Tailoring Matter? Meta-Analytic Review of Tailored Print Health Behavior Change Interventions," *Psychological Bulletin*, 133(4), 673–93.

Noble, C.H., and M.P. Mokwa (1999), "Implementing Marketing Strategies: Developing and Testing a Managerial Theory," *The Journal of Marketing*, 63(4), 57–73.

Nolan, J.M., P.W. Schultz, R.B. Cialdini, N.J. Goldstein, and V. Griskevicius (2008), "Normative Social Influence is Underdetected," *Personality and Social Psychology Bulletin*, 34(7), 913–23.

Nunnally, J.C., and I. H. Bernstein (1994), *Psychometric Theory*, New York: McGraw-Hill.

Nysveen, H., P. Pedersen, and H. Thorbjornsen (2005), "Intentions to Use Mobile Services: Antecedents and Cross-Service Comparisons," *Journal of the Academy of Marketing Science*, 33(3), 330–46.

Parry, I.W.H. (2005), "Is Pay-as-You-Drive Insurance a Better Way to Reduce Gasoline than Gasoline Taxes?," *The American Economic Review*, 95(2), 288–93.

Pauly, M.V. (1968), "The Economics of Moral Hazard: Comment," *The American Economic Review*, 58(3), 531–37.

Pavlou, P.A., and D. Gefen (2004), "Building Effective Online Marketplaces with Institution-based Trust," *Information Systems Research*, 15(1), 37–59.

Peng, W. (2009), "Design and Evaluation of a Computer Game to Promote a Healthy Diet for Young Adults," *Health Communication*, 24(2), 115–27.

Premkumar, G, and M. Potter (1995), "Adoption of Computer Aided Software Engineering (CASE) Technology: an Innovation Adoption Perspective," *ACM SIGMIS Database*, 26(2-3), 105–24.

Prochaska, J.O., and C.C. DiClemente (1992), "In Search of how People Change: Applications to Addictive Behaviors," *American Psychologist*, 47(9), 1102–14.

Qreshi, I., and D. Compeau (2009), "Assessing Between-Group Differences in Information Systems Research: a Comparison of Covariance- and Component-Based SEM," *MIS Quarterly*, 33(1), 197–214.

Quené, H., and H. van den Bergh (2004), "On Multi-level Modeling of Data from Repeated Measures Designs: A Tutorial," *Speech Communication*, 43(1-2), 103–21.

Rai, A., and M. Keil (2008), "Understanding Digital Inequality: Comparing Continued Use Behavioral Models of the Socio-Economically Advantaged and Disadvantaged," *MIS Quarterly*, 32(1), 97–126.

Rea, S.A. (1992), "Insurance Classification and Social Welfare," in *Contributions to Insurance Economics*, G. Dionne, ed., Boston: Kluwer.

Reeves, B., and C. Nass (1996), *The Media Equation - How People Treat Computers, Television, and New Media Like Real People and Places*, Stanford, USA: CSLI Publications.

Rhodes, N., and K. Pivik (2011), "Age and Gender Differences in Risky Driving: The Roles of Positive Affect and Risk Perception," *Accident Analysis & Prevention*, 43(3), 923–31.

Ringle, C., S. Wende, and S. Will (2005), "SmartPLS 2.0 (M3) Beta."

Roberts, T. (1991), "Gender and the Influence of Evaluations on Self-assessment in Achievement Settings," *Psychological Bulletin*, 109(2), 297–308.

Robinson, J.P., P.R. Shaver, and L.S. Wrightsman (1991), *Measures of Personality and Social Psychological Attitudes*, San Diego, USA: Academic Press.

Roca, J.C., J.J. Garcia, and J.J. de la Vega (2009), "The Importance of Perceived Trust, Security and Privacy in Online Trading Systems," *Information Management & Computer Security*, 17(2), 96–113.

Rogers, E.M. (1983), *Diffusion of Innovations*, New York: The Free Press.

Rothensee, M. (2008), "User Acceptance of the Intelligent Fridge: Empirical Results from a Simulation," in *Proceeding of The Internet of Things 2008*, C. Floerkemeier, M. Langheinrich, E. Fleisch, F. Mattern, and S. E. Sarma, eds., Berlin / Heidelberg: Springer, 123–39.

Rothschild, M., and J. Stiglitz (1976), "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information," *The Quarterly Journal of Economics*, 90(4), 629–49.

Santos, G.J. (2006), "Card Sort Technique as a Qualitative Substitute for Quantitative Exploratory Factor Analysis," *Corporate Communications: An International Journal*, 11(3), 288–302.

Schepers, J., and M. Wetzels (2007), "A Meta-Analysis of the Technology Acceptance Model: Investigating Subjective Norm and Moderation Effects," *Information & Management*, 44(1), 90–103.

Schloderer, M., C. Ringle, and M. Sarstedt (2009), "Einführung in die varianzbasierte Strukturgleichungsmodellierung. Grundlagen, Modellevaluation und Interaktionseffekte am Beispiel von SmartPLS," in *Theorien und Methoden der Betriebswirtschaft*, M. Schwaiger and A. Meyer, eds., München: Vahlen, 583–611.

Scholderer, J., and I. Balderjahn (2005), "PLS versus LISREL: Ein Methodenvergleich," in *Handbuch PLS-Pfadmodellierung. Methode, Anwendung, Praxisbeispiele*, F. Bliemel, A. Fassott, and J. Henseler, eds., Stuttgart: Schäffer-Poeschel, 87–98.

Schröder, A. (2006), "Prinzipien der Panelanalyse," in *Methodik der empirischen Forschung*, S. Albers, D. Klapper, U. Konradt, A. Walter, and J. Wolf, eds., Wiesbaden: Gabler.

Schultz, P.W. (1999), "Changing Behavior with Normative Feedback Interventions: A Field Experiment on Curbside Recycling," *Basic and Applied Social Psychology*, 21(1), 25–36.

Shapiro, S.P. (1987), "The Social Control of Impersonal Trust," *American Journal of Sociology*, 93(3), 623–58.

Shavell, S. (1979), "On Moral Hazard and Insurance," *The Quarterly Journal of Economics*, 93(4), 541–62.

Sheng, H., F.F.H. Nah, and K. Siau (2008), "An Experimental Study on Ubiquitous Commerce Adoption: Impact of Personalization and Privacy Concerns," *Journal of the Association for Information Systems*, 9(6), 344–76.

Shin, D.-H. (2009), "Towards an Understanding of the Consumer Acceptance of Mobile Wallet," *Computers in Human Behavior*, 25(6), 1343–54.

Shiraishi, M., Y. Washio, C. Takayama, V. Lehdonvirta, H. Kimura, and T. Nakajima (2009), "Using Individual, Social and Economic Persuasion Techniques to Reduce CO2 Emissions in a Family Setting," in *Proceedings of the 4th International Conference on Persuasive Technology*, New York: ACM.

Stevens, J.P. (2002), *Applied Multivariate Statistics for the Social Sciences*, Hillsdale, USA: Erlbaum.

Stockard, J., A.J.C. van de Kragt, and P.J. Dodge (1988), "Gender Roles and Behavior in Social Dilemmas: Are There Sex Differences in Cooperation and in its Justification?," *Social Psychology Quarterly*, 51(2), 154–63.

Stone, M. (1974), "Cross-Valedictory Choice and Assessment of Statistical Predictions," *Journal of the Royal Statistical Society, Series B*, 36(2), 111–47.

Subramanian, K., J. Lemaire, J.C. Hershey, M.V. Pauly, K. Armstrong, and D.A. Asch (1999), "Estimating Adverse Selection Costs from Genetic Testing for Breast and Ovarian Cancer: The Case of Life Insurance," *The Journal of Risk and Insurance*, 66(4), 531–50.

Sun, H., and P. Zhang (2006), "The Role of Moderating Factors in User Technology Acceptance," *International Journal of Human-Computer Studies*, 64(2), 53–78.

Taylor, T., and P.A. Todd (1995), "Understanding Information Technology Usage: A Test of Competing Models," *Information Systems Research*, 6(2), 144–76.

Tenenhaus, M., V.E. Vinzi, Y.-M. Chatelin, and C. Lauro (2005), "PLS Path Modeling," *Computational Statistics & Data Analysis*, 48(1), 159–205.

Thompson, R., D.W. Barclay, and C.A. Higgins (1995), "The Partial Least Squares Approach to Causal Modeling: Personal Computer Adoption and Use as an Illustration," *Technology Studies*, 2(2), 284–324.

Thompson, R.L., C.A. Higgins, J.M. Howell, and C. Na (1991), "Personal Computing: Toward a Conceptual Model of Utilization," *Management Information Systems*, 15(1), 125–43.

Toledo, T., O. Musicant, and T. Lotan (2008), "In-Vehicle Data Recorders for Monitoring and Feedback on Drivers' Behavior," *Transportation Research Part C: Emerging Technologies*, 16(3), 320–31.

Tormala, Z.L., P. Briñol, and R.E. Petty (2006), "When Credibility Attacks: The Reverse Impact of Source Credibility on Persuasion," *Journal of Experimental Social Psychology*, 42(5), 684–91.

Venkatesh, V. (1999), "Creation of Favorable User Perceptions: Exploring the Role of Intrinsic Motivation," *MIS Quarterly*, 23(2), pp. 239–60.

Venkatesh, V., M.G. Morris, F.D. Davis, and G.B. Davis (2003), "User Acceptance of Information Technology: Toward a Unified View," *MIS Quarterly*, 27(3), 425–78.

Venkatesh, V., M.G. Morris, and P.L. Ackerman (2000), “A Longitudinal Field Investigation of Gender Differences in Individual Technology Adoption Decision-Making Processes,” *Organizational Behavior and Human Decision Processes*, 83(1), 33–60.

Venkatesh, V., and F.D. Davis (2000), “A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies,” *Management Science*, 46(2), 186–204.

Venkatesh, V., and M.G. Morris (2000), “Why don’t Men ever Stop to Ask for Directions? Gender, Social Influence, and their Role in Technology Acceptance and Usage Behavior,” *MIS Quarterly*, 24(1), 115–40.

Venkatraman, M.P., and D.J. MacInnis (1985), “The Epistemic and Sensory Exploratory Behaviors of Hedonic and Cognitive Consumers,” *Advances in Consumer Research*, 12(1), 102–7.

Verbeke, G., and G. Molenberghs (2000), *Linear Mixed Models for Longitudinal Data*, New York: Springer.

Vickrey, W. (1968), “Automobile Accidents, Tort Law, Externalities, and Insurance: An Economist’s Critique,” *Law and Contemporary Problems*, 33(3), 464–87.

Weiss, R.E. (2005), *Modeling Longitudinal Data*, New York: Springer.

West, B.T., K.B. Welch, and A.T. Galecki (2007), *Linear Mixed Models. A Pracitcal Guide Using Statistical Software*, Boca Raton, USA: Chapman & Hall/CRC.

Westin, A.F. (1967), *Privacy and Freedom*, New York: Atheneum.

Wilcox, R.R. (2005), *Introduction to Robust Estimation and Hypothesis Testing*, Burlington, USA: Elsevier.

Yousafzai, S.Y., G.R. Foxall, and J.G. Pallister (2007), “Technology Acceptance: a Meta-Analysis of the TAM: Part 1,” *Journal of Modelling in Management*, 2(3), 251–80.

Zaichkowsky, J.L. (1985), “User Measuring the Involvement Construct,” *Journal of consumer research*, 12(3), 341–52.

Zantema, J., D. van Amelsfort, M. Bliemer, and P. Bovy (2008), “Pay-as-You-Drive Strategies: Case Study of Safety and Accessibility Effects,” *Transportation Research Record: Journal of the Transportation Research Board*, 2078(-1), 8–16.

Curriculum Vitae

Personal Details

Name	Peter Mayer
Date of Birth	1st March 1975
Place of Birth	Burghausen (Bavaria, Germany)
Nationality	German

Education

2005 - 2012	**University of St. Gallen (HSG)** Doctoral Program at the Institute of Technology Management (ITEM-HSG)
1997 - 2002	**University of Karlsruhe** Diploma in Industrial Engineering and Management (Dipl.-Wi.-Ing.)
1995 - 1997	**Technical University of Munich** Studies in Electrical Engineering

Professional Experience

2009 - present	**University of St. Gallen (HSG)** Research Associate at the Institute of Technology Management (ITEM-HSG)
2004 - 2009	**SAP** Research Associate at SAP Research
2002 - 2004	**T-Systems** Consultant
1999 - 2002	**T-Systems / debis Systemhaus** Software Developer (Part Time Employee)
1998 - 1999	**University of Karlsruhe** Student Assistant at the Institute for Computer Design and Error Tolerance