Do you really know your consumers? : analyzing the impact of consumer knowledge on use and failure evaluation of consumer electronics

DOI:
10.6100/IR658063

Document status and date:
Published: 01/01/2010

Document Version:
Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Do You Really Know Your Consumers?

Analyzing the Impact of Consumer Knowledge on Use
and Failure Evaluation of Consumer Electronics

Do You Really Know Your Consumers?

Analyzing the Impact of Consumer Knowledge on Use
and Failure Evaluation of Consumer Electronics

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de
Technische Universiteit Eindhoven, op gezag van de
rector magnificus, prof.dr.ir. C.J. van Duijn, voor een
commissie aangewezen door het College voor
Promoties in het openbaar te verdedigen
op dinsdag 9 maart 2010 om 16.00 uur

door

Jeroen Keijzers

geboren te Roosendaal en Nispen

Dit proefschrift is goedgekeurd door de promotoren:

prof.dr.ir. P.H. den Ouden
en
prof.dr.ir. J.H. Eggen

Copromotor:
dr. Y. Lu

CIP-DATA LIBRARY TECHNISCHE UNIVERSITEIT EINDHOVEN

Keijzers, J.

Do You Really Know Your Consumers – Analyzing the Impact of Consumer Knowledge on Use and Failure Evaluation of Consumer Electronics / By J. Keijzers. – Eindhoven: Technische Universiteit Eindhoven, 2010.
– Proefschrift –

ISBN 978-90-386-2170-8

NUR 964

Keywords: Product development / Consumer knowledge / Consumer complaints / Failure attribution / Consumer electronics / Product quality

Printed by: University Printing Office, Eindhoven

Cover design by: Sofie den Ouden (VissenCom)

This work has been carried out as part of the TRADER project under the responsibility of the Embedded Systems Institute. This project is partially supported by the Dutch Government under the Bsik program.

Acknowledgements

In November 2005 I started as a Ph.D. student in the sub department of Quality and Reliability Engineering at the faculty of Technology Management. Now, more than four years later, I end up completing my research at the sub department of Business Process Design at the faculty of Industrial Design. This challenging, sometimes tough, but above all interesting and rewarding journey could not have been possible without the support of many people of whom I would like to thank some in particular.

First, I would like to show gratitude to my supervisors starting with my first promotor, prof. Elke den Ouden. At the moment she was supervising my Master's project she was also working on finalizing her own Ph.D. project and now, almost six years later, she supervised me until the completion of my Ph.D. project. I would like to thank her for showing confidence in me during this process, for giving me the freedom to choose my own direction and for the numerous valuable discussions on the content of my research and on my personal development. Thank you! At the same time I met prof. Elke den Ouden as supervisor of my Master's project, I also met dr. Lu Yuan. As co-promotor of my Ph.D. project she supported me from day one with critical yet stimulating and challenging feedback on my research. She always inspired me (and still does) to look for research opportunities and to further develop myself. Words cannot express my gratitude!

I would have never started this Ph.D. project if prof. Aarnout Brombacher would not have convinced me that a doing multidisciplinary Ph.D. project with industrial partners is certainly not only about reading books and papers in a library. During the past four years he was closely involved in all the stages of this project and his feedback on my dissertation and also on my professional development is warmly appreciated. Furthermore, I would like to thank my second promotor prof. Berry Eggen for his critical view on the set-up of my experiments and for reviewing my dissertation chapter-by-chapter. I would also like to show my appreciation for the review of my dissertation by Prof. Van Gemund and Prof. Goffin. They provided me with valuable comments to improve the readability, the line of reasoning and the quality of the presentation of the empirical results in this dissertation.

I would like to thank the members of the TRADER project and the industrial and academic partners for making this Ph.D. project possible. In particular I need to thank Frans Beenker, Dave Watts, Jozef Hooman, Teun Hendriks and Roland Mathijssen from ESI and Ben Pronk from NXP. For all their invaluable help with the selection and design of the failure scenarios, I would like to acknowledge the significant support from Rob Golsteijn and Iulian Nitescu for which I would like to thank them.

During my Ph.D. research I have consulted several people for advice whom I would like to thank personally. In this context I first would like to thank Herman Hartmann for giving me a crash course on software reliability in the early stages of the project. I would also like to thank Hans van der Bij for his help in selecting and setting up the factor analysis discussed in Chapter 4. Furthermore, I would like to thank Anne Rozinat and Ton Weijters for providing advice and support on the use of process mining tools for the analysis of the usage patterns discussed in Chapter 5. I would like to thank Jun Hu for his technical support on the use of the web-based survey software used for the experiments discussed in Chapter 6 and 7. Finally, I would like to thank prof. Jean-Bernard Martens and Evan Karapanos for their help with the statistical analysis of the picture quality comparisons discussed in Chapter 7.

I would also like to thank the students Wanda Chrisiana, Jeroen Cox, Pieter Hofstra, Yvonne Kleuskens, Martin Kools and Laurie Scholten who contributed to the surveys and experiments reported in this dissertation.

Thank you to all the colleagues in the Business Process Design group at the Faculty of Industrial Design. Elke, Lu Yuan, Ilse, Hanneke, Christelle, Aylin, Renate, Aarnout, Wim, Aravindan, Kostas, Maurits, Girish, Joël, Peter and last but not certainly not least Jan: Thank you for all your help and support and for being such wonderful colleagues.

Tot slot wil ik mijn familie en vrienden bedanken voor hun steun. Pa en ma, bedankt voor alles. Jullie vertrouwen en steun hebben mij altijd gestimuleerd om het beste eruit te halen. Robbie en Maurits, ik ben vereerd dat jullie mijn paranimf willen zijn. Berry, er zijn geen woorden om jou te bedanken voor al je vertrouwen, begrip en steun tijdens deze drukke jaren. Zonder jouw steun was ik er niet aan begonnen en was het mij nooit gelukt!

Jeroen Keijzers
January 2010

Summary

Do you really know your consumers? - Analyzing the impact of consumer knowledge on use and failure evaluation of consumer electronics

The field of Consumer Electronics (CE) can be characterized by continuous technological innovation, fierce global competition, strong pressure on time-to-market, fast adoption cycles and increasingly complex business processes. In this context it is increasingly challenging for product designers and developers to provide products with unique features and excellent price / performance characteristics, as well as having to provide products that meet all the consumer's expectations. From a business perspective, research has shown that the number of consumer complaints and even product returns is increasing for complex CE (Den Ouden, 2006). Further research on the causes of these complaints showed that almost half of the complaints were due to non-technical reasons. Therefore, more insight is needed into product quality and reliability from a consumer point of view.

A literature review showed that quality and reliability methods that are currently used in product development insufficiently prevent the large variety of consumer complaints: the number of consumer complaints is rising while at the same time the root cause of these complaints is more difficult to retrace. Product failures need to be measured and analyzed from a consumer's point of view since the traditional fault-complaint propagation model fails to capture all potential sources of consumer complaints. More insight is needed into the relation between the diversity of consumers and the propagation of product development faults to these "Consumer-Perceived Failures" (CPFs).

A conceptual framework was developed to model the underlying factors related to the propagation of product development faults to consumer complaints from a consumer point of view. This framework is based on insights from human-computer interaction and consumer behavior literature and the results of an explorative experiment. Furthermore, the most commonly used consumer selection criteria for consumer tests based on demographics and/or product adoption related characteristics do not sufficiently cover differences in CPFs. The consumer characteristic "consumer knowledge" is hypothesized to have a strong impact on differences in the underlying variables of this framework. A review of relevant consumer models and consumer characteristics used in human-computer interaction and consumer behavior research shows that this construct relates to cognitive structures consumers have about a product's functioning as well as cognitive processes needed to use a complex CE product. This dissertation therefore aimed to investigate the hypothesized effect of consumer knowledge on two important variables of the conceptual framework: product usage behavior and failure attribution.

By using multiple surveys, two laboratory experiments and a web-based experiment, the following aspects of the conceptual framework were investigated in this dissertation:

- How and to what extent consumers can be differentiated on knowledge of complex CE
- The effect of consumer knowledge on differences in product usage behavior
- The effect of consumer knowledge on differences in attribution of product failures

The results of the surveys to differentiate consumers on knowledge (both core and supplemental domains) of innovative LCD televisions demonstrated the successful development and validation of measurements of both subjective and objective measurements of expertise and familiarity. It was concluded that the selection of consumer knowledge constructs as criterion for differentiating consumers for a consumer test depends on the target consumer group for a product (e.g. a very narrow homogeneous consumer group versus mass consumer markets), the type of product (e.g. passive versus active interaction) and the goal of the consumer test.

The laboratory experiment which investigated the effect of subjective expertise and objective familiarity on product usage behavior showed that higher levels of subjective expertise on both the television and computer domain result in significantly better effectiveness and efficiency and less interaction problems when performing complex product related tasks. Next, the results also showed that differences in subjective expertise stronger relate to differences in product usage behavior than those in objective familiarity. The findings of this study help product developers and designers to better understand differences in product usage behavior when consumers encounter interaction problems and can therefore help the product designers and developers to take better design decisions.

The results of both failure attribution experiments with simulated failure scenarios of picture quality failures in an LCD television showed that only objective expertise differences affect differences in consumer perception of product failures. However, although the failure attribution of consumers with higher levels of objective expertise has more dimensions and is more refined, higher levels of objective expertise on a product do not automatically result in attributions that are more in accordance with the real physical cause of the failure. This has important implications because currently used test methods often differentiate consumers only on previous experience (i.e. familiarity) with a product. The results of both studies also demonstrated that both failure cause and failure impact do not significantly affect how consumers attribute the failures.

In total it can be concluded that, when evaluating the effect of consumer diversity on fault-complaint propagation, consumer knowledge can be used to differentiate product use and failure attribution for complex CE. However, it should be noted that especially for failure attribution this effect is not consistent across different types of failures. In addition, compared to objective and subjective familiarity and subjective expertise, objective expertise has the strongest impact. In the context of fast evolving complex CE, objective expertise measurements are becoming increasingly important because familiarity or subjective expertise measurements on the (technical) functioning of currently available products can quickly become “incorrect” or “incomplete” for the next generation of products. These insights can support product designers and developers to make the right design decisions to enhance consumer satisfaction.

Samenvatting

Do you really know your consumers? - Analyzing the impact of consumer knowledge on use and failure evaluation of consumer electronics

Het vakgebied van de Consumenten Elektronica (CE) wordt gekenmerkt door doorlopende technische innovatie, door sterke wereldwijde concurrentie, door grote druk op de doorlooptijd tot marktintroductie, door de snelle aankoopcycli en door de in complexiteit toenemende bedrijfsprocessen. In deze context is het voor productontwerpers en –ontwikkelaars steeds moeilijker om producten te leveren met unieke features en een goede prijs-kwaliteitverhouding, die eveneens moeten voldoen aan de verwachtingen van consumenten. Onderzoek vanuit het bedrijfsperspectief heeft aangetoond dat het aantal klachten van consumenten en zelfs het aantal producten dat wordt geretourneerd voor complexe CE toeneemt (Den Ouden, 2006). Nader onderzoek naar de oorzaak van deze klachten laat zien dat bijna de helft van de klachten te wijten is aan niet-technische oorzaken. Daarom is meer inzicht in de productkwaliteit en –betrouwbaarheid vanuit het oogpunt van de consument noodzakelijk.

Literatuuronderzoek heeft aangetoond dat de onderzoeksmethoden voor kwaliteit en betrouwbaarheid die op dit moment voor productontwikkeling worden gebruikt in onvoldoende mate (de grote verscheidenheid aan) consumentenklachten kunnen voorkomen: het aantal klachten neemt toe terwijl tegelijkertijd de oorzaak van deze problemen moeilijker te traceren is. Productfouten moeten worden gemeten en geanalyseerd vanuit het oogpunt van de consument omdat het traditionele fout-klacht-escalatiemodel niet in staat is alle mogelijke oorzaken van consumentenklachten te ondervangen. Daarom is meer inzicht nodig in de relatie tussen de diversiteit in consumenten en de escalatie van productontwikkelingsfouten in relatie tot deze “door de Consument gePercipieerde Fouten” (CPFs).

In dit proefschrift is een conceptueel model ontwikkeld om de onderlinge factoren, gelieerd aan de escalatie van productontwikkelingsfouten tot consumentenklachten, vanuit consumentenoogpunt te modelleren. Dit model is gebaseerd op inzichten uit literatuur over mens-machine-interactie en consumentengedrag en op de resultaten van een exploratief experiment. De resultaten van dit literatuuronderzoek laten ook zien dat de meest gebruikte selectiecriteria om consumenten voor consumententests te selecteren, gebaseerd op demografische gegevens en/of productaankoopkenmerken, bovendien niet in voldoende mate de verschillen in CPFs afdekken. Het kenmerk “kennis van consumenten” heeft naar verwachting een sterke invloed op verschillen in de onderliggende variabelen van dit model. Onderzoek naar relevante consumentenmodellen en consumentenkenmerken, zoals deze bij onderzoek naar mens-machine-interactie en consumentengedrag gebruikt worden, laat

namelijk zien dat de kennis van consumenten verband houdt met de cognitieve structuren die consumenten hebben over het functioneren van een product, evenals met de cognitieve processen die nodig zijn om een complex CE product te gebruiken. Deze dissertatie had daarom als doelstelling om te onderzoeken wat de invloed is van de kennis van consumenten op twee belangrijke variabelen van het conceptueel model: productgebruiksgedrag en foutattributie. Door gebruik te maken van meerdere enquêtes, van twee laboratoriumexperimenten en van een experiment via Internet, zijn de volgende aspecten van het conceptueel model in deze dissertatie onderzocht:

- Hoe en in welke mate consumenten op basis van kennis van complexe CE onderscheiden kunnen worden.
- Het effect van de kennis van consumenten op verschillen in productgebruiksgedrag.
- Het effect van de kennis van consumenten op verschillen in foutattributie.

Uit de resultaten van de enquêtes om de gebruikers te differentiëren op basis van kennis (zowel basiskennis als aanvullende kennis) over innovatieve LCD televisies, is de succesvolle ontwikkeling en validatie van zowel subjectieve als objectieve metingen van expertise en vertrouwdheid aangetoond. Er werd geconcludeerd dat de selectie van begrippen van kennis van consumenten als criterium om consumenten voor een consumententest te onderscheiden afhankelijk is van de doelgroep voor een product (zoals een erg smalle, homogene groep consumenten versus een massa consumentenmarkt), het type product (bijvoorbeeld passieve versus actieve interactie) en het doel van de consumententest.

Het laboratoriumexperiment dat het effect van subjectieve expertise en objectieve vertrouwdheid op productgebruiksgedrag onderzocht, liet zien dat hogere niveaus van subjectieve expertise op het gebied van zowel televisies als computers resulteren in significant betere effectiviteit en efficiency en minder interactieproblemen op het moment dat complexe productgerelateerde taken worden uitgevoerd. Uit de resultaten bleek daarnaast dat de verschillen in subjectieve expertise sterker correleren met de verschillen in productgebruiksgedrag dan met de verschillen in objectieve vertrouwdheid. De resultaten van dit onderzoek zorgen ervoor dat productontwikkelaars en –ontwerpers verschillen in productgebruiksgedrag beter begrijpen als gebruikers interactieproblemen ervaren en de resultaten kunnen de productontwerpers en –ontwikkelaars aldus helpen om betere ontwerpbeslissingen te nemen.

Uit de resultaten van beide experimenten met de gesimuleerde foutscenario's voor de foutattributie betreffende de beeldkwaliteit in een LCD televisie, blijkt dat enkel verschillen in objectieve expertise invloed hebben op de wijze waarop consumenten productfouten interpreteren. Hoewel de foutattributie van consumenten met hogere objectieve expertise meer dimensies heeft en verfijnder is, hoeft een hoger objectieve expertise niveau echter niet automatisch te resulteren in attributies die meer in overeenstemming zijn met de daadwerkelijke fysieke oorzaak van de fout. Dit heeft belangrijke gevolgen, omdat de huidige testmethoden vaak differentiëren naar eerdere ervaringen (bijvoorbeeld vertrouwdheid) met een product. De resultaten van beide studies toonden ook aan dat zowel de oorzaak van de fout en de impact van de fout niet significant beïnvloeden waaraan consumenten de fout attribueren.

Alles overziend kan worden geconcludeerd dat wanneer wordt gekeken naar de invloed van verscheidenheid in consumenten op de relatie tussen fout en klacht, kennis van consumenten kan worden gebruikt om onderscheid te maken tussen productgebruik en foutattributie voor complexe CE. Er dient echter te worden opgemerkt dat vooral deze invloed op foutattributie niet consistent is wanneer wordt gekeken naar verschillende fouttypen. In aanvulling daarop kan worden gezegd dat, vergeleken met objectieve en subjectieve vertrouwdheid, objectieve expertise de grootste invloed heeft. In de context van snel veranderende, complexe CE, worden objectieve expertise metingen steeds belangrijker omdat vertrouwdheid en subjectieve expertise metingen naar het (technische) functioneren van momenteel verkrijgbare producten snel “incorrect” of “incompleet” kunnen worden wat de volgende generatie producten betreft. Deze inzichten kunnen product ontwerpers en –ontwikkelaars helpen om de juiste keuzes te maken om de consumenttevredenheid te vergroten.

Table of Contents

External Publications Related to the Dissertation

Overall

Keijzers, J. & Luyk, I.M. (2009). User perception of product failures. In R.W.M. Mathijssen (Ed.), *Trader: Reliability of High-Volume Consumer Products. (pp. 9-23).* Eindhoven, The Netherlands: Embedded Systems Institute.

Chapter 1

Keijzers, J., Den Ouden, P.H. & Brombacher, A.C. (2006). Evaluating test methods in dealing with customer perceived failures in highly innovative product development. In *Proceedings of the IEEE International Conference on Management of Innovation and Technology*, volume 2 (pp. 576–580). Singapore: IEEE.

Chapter 1 and 2 (context of research problem)

Keijzers, J., Den Ouden, P.H. & Lu, Y. (2008). The 'Double-Edged Sword' of high-feature products: An explorative study of the business impact. In *Proceedings of the 32nd Annual Product Development and Management Association (PDMA) International Research Conference*. (pp. 13-17). Orlando: PDMA.

Keijzers, J., Den Ouden, P.H. & Lu, Y. (2008). Usability benchmark study of commercially available smart phones: Cell phone type platform, PDA type Platform and PC type platform. In G.H. ter Hofte, I. Mulder (Eds.), In *Proceedings of the 10th International Conference on Human-Computer Interaction with Mobile Devices and Services, Mobile HCI 2008.* (pp. 265-272). Amsterdam: ACM.

Chapter 6

Keijzers, J., Den Ouden, P.H. & Lu, Y. (2009). Understanding consumer perception of technological product failures: An attributional approach. In *Proceedings of the 27th International Conference Extended Abstracts on Human Factors in Computing Systems*, (pp. 4057–4062). New York: ACM.

Keijzers, J., Scholten, L., Lu, Y. & Den Ouden, P.H. (2009). Scenario-based evaluation of perception of picture quality failures in LCD televisions. In R. Roy & E. Shebab (Eds.), *Proceedings of the 19th CIRP Design Conference.* (pp. 497–503). Cranfield: Cranfield University Press.

List of Abbreviations

ANOVA	Analysis Of Variance
ASQ	After-Scenario Questionnaire
CE	Consumer Electronics
CPF	Consumer-Perceived Failure
CRT	Cathode Ray Tube (television)
DTV	Digital Television
DVD	Digital Versatile Disc
ESI	Embedded Systems Institute
HCI	Human-Computer Interaction
HCCT	High Contrast Consumer Test
HD	High-Definition (television)
HDMI	High-Definition Multimedia Interface
LCD	Liquid Crystal Display (television)
MANOVA	Multivariate Analysis of Variance
MSA	Measure of Sampling Adequacy
NFF	No Failure Found
PC	Personal Computer
PDA	Personal Digital Assistant
PDP	Product Development Process
Q&R	Quality and Reliability
TRADER	Television Related Architecture and Design to Enhance Reliability
TV	Television
UI	User Interface
UPFS	User-Perceived Failure Severity
URL	Uniform Resource Locator
UTAUT	Unified Theory of Acceptance and Use of Technology
VCR	Videocassette Recorder

1 Introduction

The research presented in this dissertation deals with Quality and Reliability (Q&R) of complex high-volume Consumer Electronics (CE). This dissertation will specifically focus on the increase of the number and diversity of consumer complaints which are related to increasing uncertainty in the Product Development Process (PDP). To support effective decision making, more insight is needed into the relation between consumers and the propagation[1] of product development faults to consumer-perceived failures and consumer complaints.

First, in section 1.1, the implications of the increase of complexity of CE are discussed from a consumer, a product technological and a PDP point of view. Section 1.2 discusses the problems addressed in this dissertation. Subsequently, in section 1.3 the goal of this dissertation and ways by which this dissertation aims to contribute to this goal are presented. Since many concepts used in this dissertation have different meanings in different research contexts, in section 1.4 an overview is given of the definitions and use of the most important concepts as they are used in this dissertation. Finally, in section 1.5, the outline of the dissertation is presented.

1.1 General introduction

1.1.1 Research context

The field of CE is increasingly challenging for product design and development. Technology advances at an exponential rate, making solutions and products possible (e.g. watching television on a mobile phone) that were not feasible a decade ago (R.G. Cooper, 2001). Further fuelled by fierce global competition, CE manufacturers are integrating a growing number of new technologies to satisfy consumers' preference for high-feature products. In this context, CE is a general term referring to electronic equipment intended for everyday use by consumers. Examples of CE are MP3 players, Liquid Crystal Display (LCD) Televisions (TVs), smart phones and multimedia entertainment centers. To achieve this new functionality, the complexity of CE is increasing, both from a product internal, technological point of view, and from a product external, consumer point of view (Norman, 1998).

From a technological point of view, advances in technology result in an increasing number and diversity of features that are realized by embedded new product technologies

[1] In this context, propagation refers to how product development faults escalate to consumer complaints by going through several stages.

(Brombacher, Sander, Sonnemans & Rouvroye, 2005; Den Ouden, 2006). For example, there is a trend to use more open systems, such as smart phone operating systems, that continuously communicate with and depend on input from their environment (Siewiorek, Chillarge & Kalbarczyk, 2004). Another example is the trend to use more intelligent technologies that provide context and user dependent applications and information (Aarts & Ecarnação, 2006). Such developments combined with the consequences of Moore's law originating from the computer industry, lead to a continuous increase of software content (i.e. in terms of lines of code) in CE (Rooijmans, Aerts & Genuchten, 1996; Siewiorek et al., 2004). In fact, software has taken over many of the traditional hardware implementations in CE, making software more important for a product's Q&R. Furthermore, although hardware failures are less prominent due to effective Q&R methods (Brombacher et al., 2005, Den Ouden, 2006; Siewiorek et al., 2004), for complex systems such as CE, the number of sources for product faults due to software defects is increasing (Siewiorek et al., 2004). Due to the increasing state space of software (i.e. the collection of all possible configurations of the software), the difficulty to specify all the interactions with software and hardware from 3rd parties in all possible configurations in the consumer's usage environment and increasing pressure to reduce time-to-market, developing software with zero defects is economically not feasible (Siewiorek et al., 2004). Consequently, software in CE inherently contains flaws that can lead to various kinds of undesired product behavior varying from barely noticeable small interruptions of a function to a complete lock-up of the system (Stroucken, Seeverens, Beenker & Watts, 2005).

From a consumer point of view, these developments lead to an increase in complexity experienced by consumers during the usage of CE. First of all, research shows that, although consumers initially choose high-feature products, during product use, product usability is more important than product functionality (Rust, Thompson & Hamilton, 2006; Thompson, Hamilton & Rust, 2005). However, because of the increase of product complexity, many features of CE are often not used and the product's behavior is difficult to understand for the average consumer without having a certain level of technological expertise (A. Cooper, 1999; Han, Yun, Kwahk & Hong, 2001; Norman, 1998; Norman, 2002). Furthermore, because consumers use a variety of products and services from different manufacturers and service providers, they often are confronted with conflicting requirements and highly complex interoperability issues (Norman, 2002). In other words, these developments lead to an increase of cognitive complexity for consumers during usage of CE (A. Cooper, 1999). An example of these developments in the context of LCD TVs is shown in Example 1.1 on the next page.

In short, it is increasingly challenging for product designers and developers to provide products with both unique features with excellent price / performance characteristics and excellent product quality, which is key to product development success (R.G. Cooper, 1999; R.G. Cooper, 2005).

LCD televisions
The TV of today can be used for far more than just watching cable TV; it can be used to access the Internet, watch digital photos stored on your digital camera and connect to a personal computer (PC) to watch downloaded movie content.

From a product technological perspective these developments are the result of the shift from the analogue television of the past into a highly complex flat screen Digital Television (DTV) system with a complex software architecture (Fischer, 2004; Stroucken et al., 2005; Tekinerdogan, Sözer & Aksit, 2008). Furthermore, they must be able to interact with a digital or analogue cable signal, the Internet (wired or even wireless), set-top box, DVD player, harddisk recorder, digital camera, game console, multimedia center, VCR, PC etc.

From a consumer point of view, this also implies that consumers have more difficulties in understanding and using televisions with far more advanced menu options, cable connectors etc., as, for example, shown by Darnell (2008).

Example 1.1 Example of increasing complexity of LCD TVs.

Interestingly, from a business perspective, research has shown that for complex CE the number of consumer complaints and even product returns is increasing (Den Ouden, 2006). This is shown in Figure 1.1. This increase in complaints not only results in more costs for complaint handling at customer service centers and helpdesks, but also has a negative effect on consumer satisfaction, word of mouth, and even repurchase intention (Day & Landon, 1977).

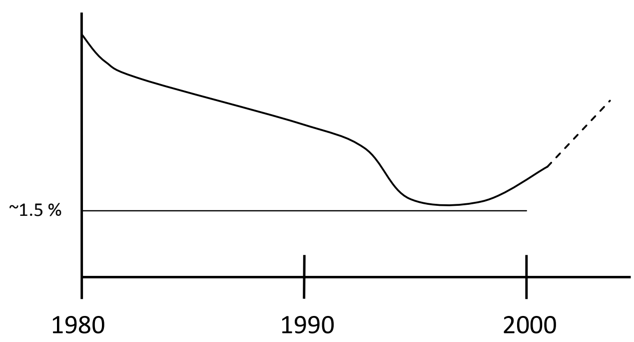


Figure 1.1 Average percentage of consumer complaints on new CE products relative to the number of products sold worldwide (Den Ouden, 2006).

More importantly, analysis of these complaints shows that this increase in complaints is not due to hardware failures (i.e. not meeting explicit product specifications), but to problems both within the product's capabilities (e.g. problems with ease of use and learning or understanding the product) and beyond the product's capabilities (i.e. not meeting consumer expectations) (Den Ouden, 2006; Koca & Brombacher, 2008). An example of how the number and diversity of consumer complaints on complex CE are increasing and what the potential business impact is, is the development of smart phones (i.e. complex, high-feature mobile communication products). These developments are discussed in Example 1.2.

Consumer complaints on smart phones

A study of mobile device returns in the United Kingdom showed that one in seven cell phones was returned as faulty within the first year of purchase (Overton, 2006). Of these returns, about 63% had no hardware or software fault but the reported problems related to usability, a mismatch with the consumer's expectations, or issues relating to the configuration of the device. Another survey in 2007 in the United States showed that 29 % of the cell phone users experienced a product failure in the past 12 months of product use (Horrigan, 2008).

A specific example of how product complexity of smart phones can lead to consumer complaints of which the root cause is difficult to determine, is the introduction of the iPhone in the Netherlands in 2008. Since its market introduction consumers report problems with the quality of the network coverage and subsequently blaim the network provider (Van Dijk, 2008). However, the network provider and other sources claim that either the product's software or a chipset from a third party manufacturer are to blaim (Krazit, 2008; Van Dijk, 2008). More recently, following consumer complaints on usability problems and software failures of the recently introduced Blackberry Storm, the manufacturer announced that due to time-to-market pressure, product failures are part of the new reality of making complex cellphones (Sharma & Silver, 2009).

Example 1.2 Example of consumer complaints in the smart phone industry.

Moreover, studies reveal that the causes of most of the product development faults associated with these complaints can be traced back to decisions made during the early phases of the PDP (Den Ouden, 2006; Koca & Brombacher, 2008). Effective decision making in the PDP of CE is increasingly difficult in a market characterized by continuous technological innovation, fierce global competition, strong pressure on time-to-market and fast adoption cycles, and increasingly complex business processes (Brombacher et al., 2005; Den Ouden, 2006). Consequently, more in-depth understanding of consumer complaints is required from both the product complexity and the consumer point of view. This dissertation will mainly focus on the consumer point of view.

1.1.2 Project context

The research discussed in this dissertation has been carried out as part of the TRADER project managed by the Embedded Systems Institute (ESI). This project is sponsored by the Dutch Ministry of Economic Affairs under the BSIK program and is carried out by a consortium of industrial and academic partners (Stroucken et al., 2005).

The previous section illustrated developments in the CE industry that led to an increase of product complexity from both a product technological as well as a consumer point of view. In this context, the TRADER project specifically focuses on broader reliability issues related to the explosive growth of software content of embedded systems in CE. Given the increasing level of product complexity, shifting error sources and strong pressure on time-to-market, zero defect software is not (economically) feasible (Siewiorek et al., 2004; Stroucken et al., 2005). The main objective of TRADER is therefore the development of methods and tools for ensuring reliability of CE resulting in the minimization of product failures that are exposed to the consumer (Stroucken et al., 2005). Within this main objective, the project focuses on Digital Television (DTV) systems as an application domain.

In general, the TRADER project aims to address the issues above by (Stroucken et al., 2005):

- Developing system architectural methods and tools for designing reliable embedded systems.
- Providing software implementation techniques for failure mode detection, failure localization and failure recovery.
- Developing a consumer-centered approach to identify and assess product failures from a consumer perspective.

The research presented in this dissertation is primarily concerned with the last research topic, consumer-centered design for reliability. It focuses on including the consumer perspective and actual consumers to identify and minimize the impact of the most important product failures.

Finally, it is important to note that this research work has been carried out partly in parallel with the research work of De Visser (2008), which also has been part of the TRADER project. Although both projects dealt with consumer-centered design for reliability, each project focused on different aspects of the identification and analysis of product failures. How these projects relate to each other, and how the research context is translated into a definition of the specific research problem addressed in this dissertation, will be discussed in the following sections.

1.2 Problem definition

Section 1.1.1 illustrated that the increase of consumer complaints on CE in an industrial context can be traced back to increasing uncertainty in decision making in PDPs of increasingly complex products. According to Mullins and Sutherland (1998), manufacturers in rapidly changing markets such as the CE industry are confronted by market, product technology and industrial chain related uncertainties, which have to be effectively managed during the PDP. Previous research showed that the existing approaches for managing product Q&R are not sufficient in the changing business context of CE as they lack consumer orientation (De Visser, 2008) and do not cover the increasing market and product technology uncertainty (Brombacher et al., 2005; Den Ouden, 2006). Additionally, the results of a literature review (presented in a separate study[2]) show that even currently used consumer test

[2] This study is published in: "Keijzers, J., Den Ouden, P.H. & Brombacher, A.C. (2006). Evaluating test methods in dealing with customer perceived failures in highly innovative product development. *In Proceedings of the IEEE International Conference on Management of Innovation and Technology, volume 2* (pp. 576–580). Singapore: IEEE".

methods do not provide sufficiently rich information on how diverse consumer groups experience product failures. In short, currently used methods do not fully cover the variability of the root causes of consumer complaints: the number of consumer complaints is rising while at the same time the root cause of these complaints is more difficult to retrace. As a result, there is a lack of understanding of product failures and subsequent consumer complaints from a consumer point of view.

To understand why this insight is currently lacking, this section further discusses the uncertainties associated with the increase of the number of consumer complaints from both the consumer and the product developer perspective. First of all, consumers have become far more demanding, more fragmented, and less predictable than they used to be. While CE used to have a single functionality and were developed for local markets, they are now becoming increasingly multifunctional, flexible and adaptive and are developed for global mass consumer markets. Furthermore, products move faster through their adoption cycles (Den Ouden, 2006), as shown in Figure 1.2. From this figure it can be seen that approximately the eighth generation of VCRs reached the late majority adopter group while already the third generation DVD recorders reached this adopter group. Consequently, it becomes far less feasible to define homogenous target consumer groups with a certain use profile with a high level of certainty (De Marez & Verleye, 2004; Grudin, 1991; Kujala & Kauppinen, 2006) compared to the development of one tailored product for a very narrow adopter group in the past (e.g. the first computer systems). Combined with the increase in cognitive complexity as discussed in section 1.1.1, the behavior of the product in the field becomes far less predictable.

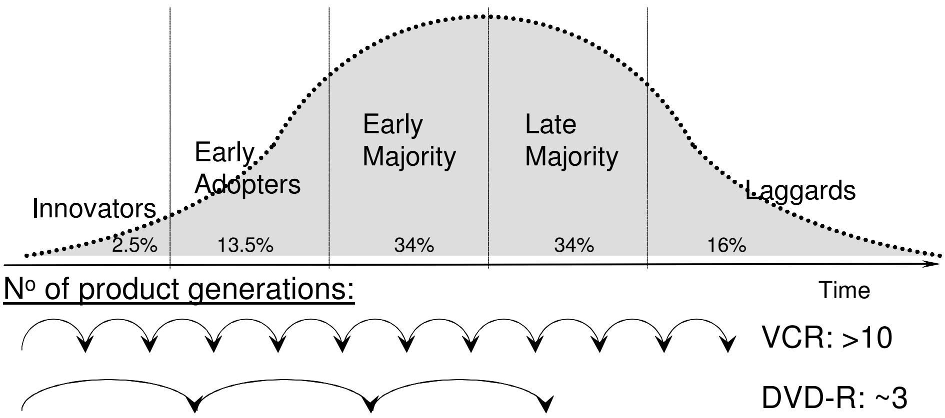


Figure 1.2 Reduced time to commodity in the product adoption cycle (Den Ouden, 2006).

Secondly, product designers have difficulty predicting and preventing consumer complaints for these large and diverse consumer groups. Research by De Visser (2008, chapter 4) shows that in practice product designers have difficulties predicting the level of dissatisfaction that consumers experience when confronted with a product failure. As discussed by A. Cooper (1999, p. 17), Norman (1998, p. 155) and Hasdoğan (1996), product designers and developers often use themselves as "target customer" and therefore as the frame of reference during product development. This implies they do not take the "normal" user of the product into account. As an example, A. Cooper (1999) discussed that software developers have a very difficult time making products easy to use for consumers who do not have the same level of knowledge on software; they often assume the consumer has a considerable (often implicit) amount of knowledge that the real consumer may lack. Consequently, Den Ouden (2006, p. 58) argues that product developers need more insight in differences among consumer groups to increase the coverage of the current reliability testing program and to be able to prevent consumer complaints before a new product enters the market.

However, research shows that there is another side to this problem. Product development faults do not always lead to consumer complaints and, vice versa, consumer complaints cannot always be (directly) attributed to faults made during product development (De Visser, 2008; Den Ouden, 2006). Research in the field of information systems shows that product development faults and their activation in the form of product errors often do not lead to visible product failures and thus consumer complaints (Aviezinis, Laprie, Randell & Landwehr, 2004). On the other hand, product behavior within specifications may be totally unacceptable for some consumers because it simply does not meet their expectations (Den Ouden, 2006; Siewiorek et al., 2004). For example, a consumer can perceive that a DVD player is malfunctioning because it does not recognize a certain DVD while it could be a part of the product's specifications not to play a dirty or damaged DVD because it would result in a decreased picture quality of the movie. Consequently, the relation between product development faults and consumer complaints is not fully understood. Product development faults are only important when they are triggered during product use, perceived as a failure and result in consumer dissatisfaction.

To capture all potential sources of consumer complaints, a broader definition of Q&R problems than only the by the product developer "acknowledged" product development faults is therefore required: *Consumer-Perceived Failures* (CPFs). In this context, a CPF refers to all situations in which the consumer *perceives* that something is actively wrong with the product which s/he may decide to report to the manufacturer and/or other parties involved (e.g. a service provider). This implies that a CPF might be due to one, or an interaction of two or all, of the following sources:

- <u>Product development fault</u>: hardware or software faults or flawed interaction between components and/or services of different parties involved
- <u>Product usage environment</u>: both the social and usage context of product use.
- <u>Consumer</u>: the consumers' own actions or perception that something is wrong (while the product is meeting the product specifications).

Summarizing, in this section it was shown that existing approaches for managing product Q&R do not cover uncertainties associated with the increase of consumer complaints for complex CE. The problem is that there is a lack of consumer insight with respect to the relation between the heterogeneous target consumer groups and the propagation of product development faults to CPFs and consumer complaints. Consequently, as discussed by Den Ouden (2006, p. 37), to be able to capture all (potential) reasons for dissatisfaction and product returns, product designers need a better understanding of the consumer experience with respect to all phases of their interaction with complex CE.

1.3 Aim of the dissertation

This dissertation aims to gain more insight into the relation between the diversity of consumers and the propagation of product development faults to CPFs and subsequent potential consumer complaints for complex CE. This insight can be used twofold:

1. To better account for the heterogeneity of the target consumer groups and to better account for the consumer's perception of product failures to improve the input for, and measurements prescribed by, currently used methods and tools to manage product Q&R.
2. To support design decisions in the PDP of CE to help prevent potential CPFs before a new product enters the market.

In the project context discussed in section 1.1.2 these insights are valuable to support product developers since they do not know how consumers will respond to software reliability improvements in the application domain. Many consumers do not know how a TV technically functions and simply respond to the observable behavior of the TV. A zero-defect product will not be feasible, but a reliable "TV-of-the-future" from a consumer perspective will be required.

This dissertation intends to contribute to this goal in three steps. First, the dissertation aims to investigate how consumer diversity and its effect on the propagation of product development faults to CPFs and subsequent consumer complaints can be modeled. In Chapter 2 it will be shown that the classical Q&R fault-complaint propagation model fails to capture all potential reasons for consumer complaints. Insights from a consumer complaint model from consumer behavior literature will be used to model the fault-complaint propagation from a consumer point of view. Subsequently, it will be shown that the currently used consumer segmentation criteria do not sufficiently cover differences in CPFs. Based on insights from Human-Computer Interaction (HCI) and consumer behavior research, in this dissertation consumers will therefore be differentiated on multiple dimensions of a single consumer characteristic that affects the consumers' understanding of complex CE: "consumer knowledge". Research by Alba and Hutchinson (1987) showed that consumer knowledge relates to both the cognitive structures consumers have (e.g. beliefs about a product's functioning) as well as the cognitive processes to be able to perform product-related tasks successfully. As such, differences in the level of consumer knowledge on complex CE will be used in this dissertation to gain more insight into the occurrence of CPFs.

Second, the dissertation will provide a conceptual framework to better understand the underlying factors related to the propagation of product development faults to consumer complaints. In Chapter 3, two important mediating variables in this framework, product usage behavior and failure attribution, will be further investigated. Product usage behavior will be measured in terms of usability measurements as well as in terms of product usage patterns. Failure attribution will be used as a measurement of the consumer's perception of a product failure cause. Oliver (1996) and Folkes (1984) have shown that failure attribution significantly influences various post-purchase behaviors such as consumer dissatisfaction and complaining behavior.

Third, the dissertation partially validates this conceptual framework by investigating how differences in consumer knowledge affect the two selected mediating variables. By relating consumer knowledge to differences in product usage behavior and failure attribution, this dissertation aims to contribute to the understanding of how consumer diversity affects the propagation of product development faults to consumer complaints.

Finally, it is important to discuss how the goals of the research presented in this dissertation relate to the goals of the research project conducted by De Visser (2008) in the same project context. It was previously discussed in section 1.1.2 that both research projects focus on consumer-centered design for reliability. In this context, the research by De Visser (2008) aimed to provide an overall high-level framework for product designers to assess the impact of potential quality problems on consumer dissatisfaction. Both a product technology (i.e. failure characteristics) and a user's point of view (i.e. user characteristics and use conditions) are integrated in this framework to provide designers a user-centered assessment of perceived failure severity. This dissertation does not focus on the level of product failure impact assessment but aims to provide insight into how certain consumer characteristics result in different CPFs which could eventually result in differences in perceived failure severity. As such both research projects are complementary and aim to support the decision making process during the PDP of complex CE by providing a consumer-focused approach.

1.4 Definition of concepts

The previous sections illustrated that the topic of this dissertation covers multiple disciplines, including Technology Management, Industrial Design, Marketing, Information Sciences and Psychology. Since many concepts used in this dissertation have a different meaning in different research contexts, this section presents an overview of the definitions and use of the most important concepts as they are used in this dissertation. A formal definition of all the concepts used in this dissertation can be found in the glossary.

First of all, because this dissertation deals with the perception of product failures in CE from the perspective of individual persons, throughout the dissertation the generic term "consumer" will be used to refer to a product's (intended or actual) user, buyer (usually referred to as "customer") or any other individual or group of interest for product development (PDMA NPD Glossary, 2009).

Furthermore, similar to the research presented in the dissertation by De Visser (2008), in this dissertation an extended quality definition is used in which quality refers to "the collection of attributes, which when present in a product, means a product has conformed to or exceeded consumer expectations" (adapted from PDMA (2009)). In this context, all situations in which a consumer perceives that something is actively wrong with the product, i.e. the product does not meet the consumer's expectations, will be referred to as a "consumer-perceived failure". As such, a CPF can originate from the product's manufacturer(s), the consumer, the environment of product use or interaction between these variables. Unless stated otherwise, no further distinction is made between hard and soft failures or reliability problems (Brombacher et al., 2005), usability or utility problems (Nielsen, 1993) or any other differentiation from the research domains involved.

Finally, the research presented in this dissertation is predominantly focused on CE with a high degree of product complexity from both a consumer and technological complexity point of view. This includes high-feature products such as smart phones, multimedia entertainment centers, game consoles etc. and excludes for example CE such as a simple alarm clock or a coffee machine.

1.5 Overview of the dissertation

In this section, an overview is given of the content and structure of the dissertation. First, in Chapter 2, the results of a literature review will be presented, which investigated the different stages of, and influencing factors on, the propagation of product development faults to CPFs and consumer complaints. These results are used to formulate a conceptual model that incorporates the consumer perspective on product failures. Subsequently, based on a review of methods and consumer segmentation criteria used to involve consumers in the PDP, it will be shown that these methods do not cover the variability of CPFs as defined in the conceptual model. This chapter therefore concludes with the argumentation for a need to differentiate consumers on deeper level characteristics instead of consumer profiles to investigate the consumer's perception of product failures. The results of a literature review on relevant characteristics will be discussed which results in an explicit choice to focus on "consumer knowledge" as main differentiator of consumers in the remainder of the dissertation.

Based on the results of the literature review discussed in Chapter 2, Chapter 3 starts with an explorative experiment to investigate the relation between one dimension of consumer knowledge (i.e. familiarity) and the propagation of an implemented product fault to CPFs. Subsequently, the insights from this experiment are used to further investigate and define two important mediating variables which affect CPFs: product usage behavior and failure attribution. On the basis of these results, this chapter concludes with a conceptual research framework, formulation of the research questions and an overview of the research approach.

In Chapter 4, the set-up and results of a survey, which was used to investigate whether consumers can be differentiated on knowledge of complex CE, will be presented. Subsequently, this differentiation is used in Chapter 5 to investigate in a laboratory experiment how consumer knowledge differences affect product usage behavior when consumers are asked to perform complex product tasks.

In Chapter 6 and 7 the hypothesized effect of consumer knowledge on failure attribution is investigated for different types of product faults in subsequently a web-based and a laboratory experiment.

Finally, in Chapter 8, the most important findings of this research are summarized and main conclusions are drawn. Furthermore, theoretical and practical implications are discussed and directions for future research are given.

2 Fault-complaint propagation from a consumer perspective

As discussed in the previous chapter, there is a lack of insight into the relation between the diversity of consumers and the propagation of product development faults to consumer complaints. This chapter presents the results of a literature review to gain more insight into this relation from a consumer point of view.

In section 2.1 a conceptual model is developed to give insight into how consumer diversity affects the different stages of the propagation of product development faults to consumer complaints. This section concludes with a discussion leading to the initial research focus. Subsequently, section 2.2 discusses different ways in which consumer diversity can be modeled to investigate differences in consumer-product interaction problems and CPFs. This section concludes with an argumentation to further focus on cognitive consumer models. Finally, in section 2.3 related research on consumer differentiation on cognitive models is discussed. This section concludes with a further narrowed down research focus.

2.1 Understanding the propagation of product development faults to consumer complaints

To better understand how diversity of consumers affects consumer complaints, this section discusses the positioning of the concept of CPFs and its antecedents in the propagation of product development faults to consumer complaints. In literature different approaches can be found which already partially address this propagation. This section reviews a model from a Q&R perspective and from a consumer behavior perspective. The gaps found in both models will be addressed in a revised conceptual fault-complaint propagation model from a consumer perspective.

2.1.1 Consumer complaints on complex consumer electronics

Traditionally, consumer complaints on and returns of CE are logged at customer call centers and service centers respectively, and analyzed to improve subsequent product generations (Petkova, 2003). However, as shown in Figure 2.1, research by Brombacher et al. (2005) has demonstrated that of an increasing percentage of these complaints, the root cause cannot be determined (i.e. so called “No-Failure-Found” (NFF)).

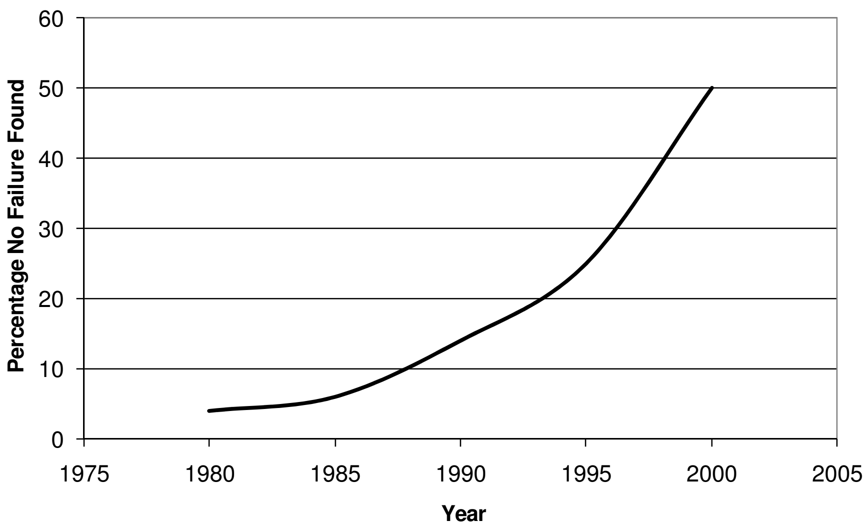


Figure 2.1 Percentage No-Failure-Found in modern high-tech, high-volume consumer electronics (Brombacher et al., 2005)

According to a recent study performed by Accenture in 2007 (Steger, Sprague & Douthit, 2007), in the USA alone approximately 13.8 billion USD is spent in the CE industry on analyzing and processing product returns. According to their study, NFF contributes to 20% of these costs. This is an important problem for CE manufacturers, and even more important when considering trends in warranty coverage which nowadays allow consumers to return their products when the product simply does not meet their expectations (Berden, Brombacher and Sander (2000)). Furthermore, as shown in Figure 2.2, complaints are only one of many ways via which consumers express dissatisfaction with a product (Day & Landon, 1977). In other words, from a manufacturer point of view, complaints are only the “tip of the iceberg” and possibly refer to many more “hidden” problems. Although not directly visible, private action as depicted in Figure 2.2 can have a significant effect on cost of non-quality in the longer run.

To be able to gain more insight into the relation between consumers and the propagation of product development faults to consumer complaints, the first step is to understand what consumers complaints are about. As discussed in Chapter 1, the increase of complaints on complex CE is not due to the product not meeting specifications alone, but due to problems both within and beyond the product’s capabilities (Koca & Brombacher, 2008). Further analysis of consumer complaints on CE in several case studies indicates that the percentage of complaints related to such problems is more than 50% (Den Ouden, 2006; Koca & Brombacher, 2008, Overton, 2006). Examples of such problems include problems with the installation and configuration of a device, connectivity problems (compatibility) with other products, not being able to understand the User Interface (UI), manual or product feedback messages etc.

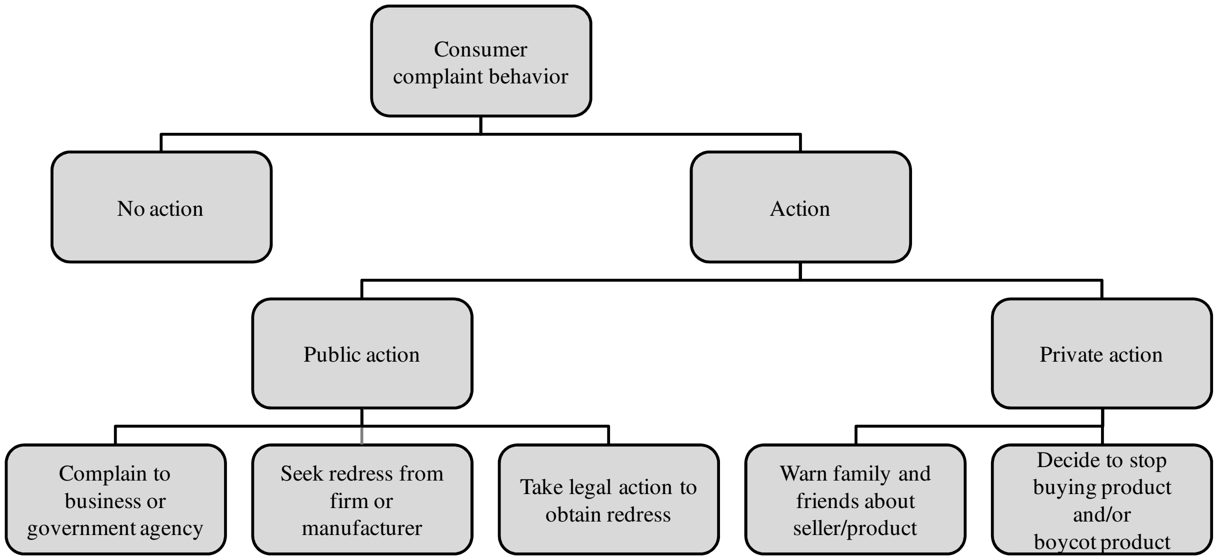


Figure 2.2 Classification of consumer complaint behavior (Day & Landon, 1977)

Research by Petkova (2003) and De Visser (2008) shows that because service centers are strongly logistically oriented to keep the service costs at a minimum, the currently used field feedback mechanisms do not provide sufficient information to be suitable for the identification of the root causes of this wide spectrum of consumer complaints. The data logged at these service centers lack information on the consumer and the context in which the problem occurred, resulting in an increase of NFF (De Visser, 2008; Koca, Karapanos & Brombacher, 2009). Besides service and helpdesk related feedback, manufacturers can also collect consumer feedback on problems via the Internet through web-based helpdesks or forums or via the seller of the product (Den Ouden, 2006; Koca, Karapanos et al., 2009). Although there are indications that this feedback is more suitable for root cause analysis (Den Ouden, 2006), it is currently still questionable whether these sources provide reliable and complete information on all potential reasons for consumer complaints. For example, it is likely that not all consumer groups use the Internet to give feedback. Moreover, research by Den Ouden (2006, chapter 5) reveals that even technical product failures are difficult to analyze and classify due to lack of contextual information on the root cause of a complaint.

Summarizing, although there is ongoing research to improve the use of field feedback to better diagnose consumer complaints and subsequently use that information in the PDP[3], it does not provide enough information to understand when product failures are triggered, perceived and reported by the consumer (see also De Visser (2008, chapter 2)). The next step is to investigate whether there are models in literature that give more insight into the context in which consumer complaints arise and how CPFs would fit in such a model. In the

[3] More information on research conducted on these topics can be found in the project descriptions of the IOP ‘Managing soft reliability’ project (Senternovem, 2005) and the IOP ‘Data fusion’ project (Senternovem, 2008). An overview of the first results of the ‘Managing soft reliability’ project can be found in Koca, Funk et al. (2009).

following sections, models from the Q&R field and from the consumer behavior field related to the fault-complaint propagation are discussed.

2.1.2 Fault-complaint propagation from a quality and reliability perspective

Derived from literature in Information Science, the propagation of faults to potential consumer complaints from a Q&R perspective consists of four phases as is shown in Figure 2.3.

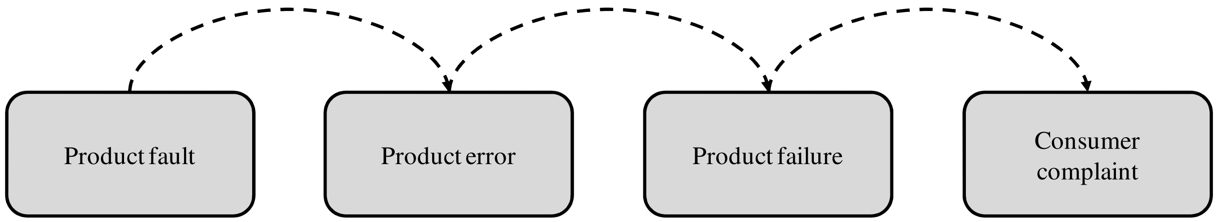


Figure 2.3 Q&R perspective on the propagation of product faults to consumer complaints (Aviezinis et al., 2004; De Visser, 2008)

This classical Q&R model, which originates from more than 20 years of research on dependable computing and fault tolerance (Aviezienis et al., 2004; Laprie, 1985), depicts a product failure as the main prerequisite for a consumer complaint. In this context a product failure is defined as (Aviezinis et al., 2004): "an event that occurs when the delivered service deviates from the correct service". Aviezienis et al. (2004) further state that this event occurs because the service deviates from the functional specification or because the specification did not adequately describe the system function. The deviation of the external state of the system from the correct service state is called a product error. Please note that from a system dependability point of view, the term "system" can also refer to an internal system of which many together form a larger system (i.e. product) from a consumer point of view. Finally, the hypothesized cause of a product error is called a product fault (Aviezinis et al., 2004). In the Q&R model dashed arrows are used to indicate that each causal relation between the elements of the model implies that one or more occurrences of a cause could potentially but not necessarily lead to the occurrence of its effect. For example, a product error in sub system A can directly lead to a product failure while a product error in sub system B, only together with the occurrence of product error in sub system C and D leads to a product failure. Furthermore, a product error in sub system E could be such that it never leads to a product failure.

Although this model originates from Information Science literature and is used to model the propagation of hardware and software faults (Aviezinis et al., 2004; Siewiorek et al., 2004), it can also be used to model the propagation of other types of potential faults in CE (De Visser, 2008, chapter 1). Given the goal of this research project and the broad definition of a product failure in Chapter 1, no further distinction will be made between product faults and product errors and all hypothesized causes of product failures will be referred to as *(product* *development) faults*. In other words, in terms of the variables shown in Figure 2.3, this dissertation does not further distinguish between *product* faults, errors and failures.

Given the research and project context discussed in section 1.1, this dissertation will generally focus on two different types of product development faults: faults related to the interaction design and faults related to the product's software (Brombacher et al., 2005; Koca & Brombacher, 2008; Siewiorek et al., 2004)[4]. Interaction design faults refer in this context to all product development mistakes resulting in problems where the product meets its specifications but the consumer cannot use, find or understand a product's functionality (Koca and Brombacher, 2008). For example, these faults relate to a bad design of the UI, manual and even because too much or the wrong functionalities are included in the product. Summarizing, this Q&R model helps to explain how mistakes made during product development can potentially result in situations in which the product is no longer functioning according to its specifications.

However, it does not fully capture all potential reasons for consumer complaints from a consumer point of view because this model is intended to give insight in the propagation of faults to consumer complaints from a technical Q&R perspective. First of all, from a Software Engineering point of view Siewiorek et al. (2004) and Chillarge (1996) discuss that system availability and software failures should be defined and analyzed from a consumer *perception* point of view since a consumer to a large extent determines whether a failure has occurred or not. In other words, a product failure from a technical, Q&R perspective is not necessarily a CPF. To understand why the concept of perception is important for understanding reasons for consumer complaints, first the concept of perception needs to be defined. According to Smith, Nolen-Hoeksma, Fredrickson and Loftus (2003, p. 190) "perception involves the translation of information acquired by our senses into a meaningful experience". A key distinction in perception is between bottom-up and top-down processes. Bottom-up processes are driven solely by input, raw sensory data while top-down processes are driven by a person's knowledge, experience, attention and expectations (Smith et al., 2003). Consequently, since not every consumer is an expert on complex CE and often does not fully understand a product's functioning (see for example A. Cooper (1999) and Norman (1998)) and since the same product is used in different usage environments, this results in differences in problem solving behavior when encountering a failure and even differences in perception of what is a failure and what is not. In other words, a fault is only a problem when the consumer perceives it as a problem. Therefore, all three antecedents of CPFs (i.e. product development faults, the consumer and the environment) need to be captured in a model to fully understand the propagation of faults to consumer complaints and to give a product designer better insight to better predict consumer complaints.

[4] Koca and Brombacher (2008) also distinguish faults related to hardware, manufacturing, back-end marketing and service but these are out of the scope of the Trader project as described in section 1.1.2.

Before addressing this gap in a revised model in section 2.1.4, another drawback of the model shown in Figure 2.3 needs to be addressed. Recent research in HCI shows that simple error metrics fail to capture the consequences consumers experience while encountering product failures (Feng & Sears, 2009). A product failure as defined in the Q&R model does not automatically result in a consumer complaint. From a psychological point of view, complaints are defined by Kowalski (1996) as "expressions of dissatisfaction, whether subjectively experienced or not, for the purpose of venting emotions or achieving intrapsychic goals, interpersonal goals or both". In other words, a deviation from correct service and even a perceived deviation of correct service do not automatically lead to a consumer complaint. To better understand the relation between CPFs, dissatisfaction and consumer complaints, a failure-complaint propagation model from a consumer behavioral perspective will be discussed in the following section.

2.1.3 Failure – complaint propagation from a consumer behavior perspective

Research in the field of consumer (complaint) behavior has shown that there are many mediating and moderating factors between the outcomes of a consumption experience (both positive and negative, such as a product failure) and the resulting levels of (dis)satisfaction and related potential complaints (Broadbridge & Marshall, 1995; Fournier & Mick, 1999; Oliver, 1996). Derived from a basic satisfaction model discussed by Oliver (1996), a theoretical consumer behavioral perspective on the propagation of performance outcomes to consumer complaints is shown in Figure 2.4. Please note that the consumer behavior model only depicts dissatisfaction since the goal of this research is to investigate consumer complaints.

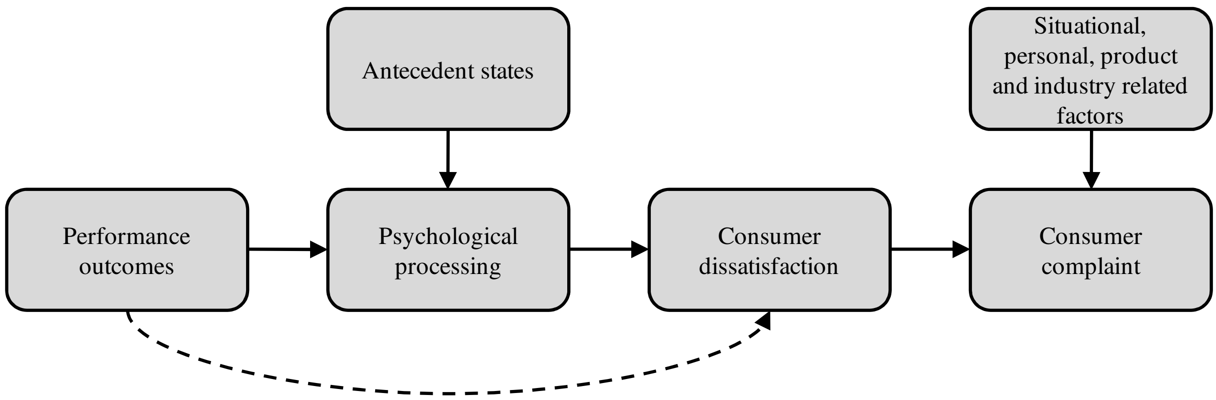


Figure 2.4 Consumer behavioral perspective on consumer complaints (derived from Oliver (1996, chapter 2))

Before discussing how and to what extent the consumer behavior model gives better insight into the propagation of faults to consumer complaints, first the concepts and functioning of this model need to be explained. In this model a performance outcome is defined as (Oliver, 1996, p. 28): "the perceived amount of product or service attribute received, usually reported on an objective scale by good and bad levels of performance”. In other words, this model takes the consumer perception of an outcome as a starting point. Subsequently, the consumer’s psychological processing mediates the impact of these performance outcomes on (dis)satisfaction judgments (Oliver, 1996, p. 40)[5]. In some instances a performance outcome can directly result in a satisfaction judgment without psychological processing, as indicated by the dashed arrow. Furthermore, the model also incorporates antecedent states (e.g. expectations or prior experience) as possible moderating variables of the psychological processing.

According to this model the emerging dissatisfaction judgment precedes a consumer complaint. In this context, dissatisfaction is defined as (Oliver, 1996, p. 28): “a judgment that a product or service feature, or the product or service itself, provided (or is providing) an unpleasant level of consumption related fulfillment, including levels of under- or overfulfilment”. Although dissatisfaction is a prerequisite for consumer complaints, research has shown that the majority of complaining behavior does not originate from simple dissatisfaction (Broadbridge & Marshall, 1995; Oliver, 1996; Tronvoll, 2007). A literature review by Tronvoll (2007) has shown that situational (e.g. moment and situation in which dissatisfaction occurs), personal (e.g. personality and emotional factors), product (e.g. durables vs. non-durables) and industry related factors (e.g. channels via which a complaint can be filed) need to be taken into account when predicting consumer complaint behavior. Although research by Broadbridge and Marshall (1995) has shown that the percentage of no-action for durables such as CE is significantly lower than for other products, these factors need to be taken into account when assessing the propagation of faults to consumer complaints. Moreover, there are indications that stand alone problems, for example related to usability, should be evaluated in the long-term context of user experience which in turn results in more meaningful (dis)satisfaction judgments (for example to assess repurchase intention etc.) (Desmet & Hekkert, 2007; Karapanos, Zimmerman, Forlizzi & Martens, 2009). This is beyond the scope of this dissertation but nevertheless indicates the complexity of consumer (dis)satisfaction judgments.

Summarizing, the consumer behavior model presented in Figure 2.4 provides insight into how, from a consumer point of view, perceived performance outcomes can potentially result in consumer complaints. Consequently, this model partially addresses the gap between failures and complaints in the Q&R model shown in Figure 2.3 in the previous section. However, since this model is from a pure consumer behavioral point of view, it does not fully address the antecedents of CPFs (i.e. consumers, the environment and product development faults) and it does not address the relation between product failures from a Q&R perspective, CPFs and perceived performance outcomes. In the following section, elements of both models will

[5] Please note that there are many factors underlying the consumer’s psychological processing of performance outcomes in satisfaction judgments (see Oliver, 1996). However, for the exploratory purpose of this section, these factors will not be further elaborated upon.

be used to develop a revised model to give insight into the relation between consumers and the propagation of faults to consumer complaints.

2.1.4 Conceptual research model

The purpose of section 2.1 is to discuss the positioning of the concept of CPF in the propagation of faults to consumer complaints. From section 2.1.1, it can be deducted that CPFs are very context and consumer dependent, which explains why it is difficult for traditional cost-optimized field feedback channels to retrieve the root cause of consumer complaints and product returns. Subsequently, in section 2.1.2 and section 2.1.3 two fault-complaint propagation models were discussed, one from a traditional Q&R perspective and one from a consumer behavior perspective. Although both models give valuable insight into how potential mistakes in the PDP can eventually lead to a consumer complaint, a new conceptual model needs to be formulated because both individual models; 1) do not address the whole propagation chain, and 2) do not explicitly cover the concept of CPFs and all of its antecedents.

As discussed in the previous sections, a product failure as defined from a Q&R perspective does not automatically lead to a CPF. To better understand how to incorporate the concept of CPFs in this propagation, it is important, as discussed in section 1.2, to consider that faults are only important when they are triggered during product use, perceived as a failure and result in consumer dissatisfaction. Consequently, product use needs to be included in the model when considering that faults can only result in CPFs when they occur during consumer-product interaction. In other words, a fault can become a consumer-product interaction problem (from now on referred to as interaction problems) when this fault is triggered during the use of the product and limits the consumer in achieving his/her goals. Differentiation between interaction problems and CPFs is important because it denotes the difference between something objectively going wrong during consumer-product interaction (e.g. a bad picture quality) and the subsequent *perception* by the consumer that this is a failure (e.g. perceived failure of the TV or the cable signal provider). In other situations an interaction problem could be solved by the consumer before it is perceived as a product failure.

Furthermore, Fournier and Mick (1999) and Rooden and Kanis (2005) discuss that it is important to consider that product usage and product failures only have a meaning when they are studied with the unpredictable variability of real consumers in realistic product usage environments. To address these gaps, the propagation model needs to incorporate both the environment and the consumer as possible antecedents of interaction problems and subsequent CPFs. Besides originating from faults during product development, the functioning and interpretation of the functioning of complex products such as CE is always dependent on the functioning of other technologies, services and infrastructure in their environment, on the consumer him/herself and on other people affected by the product use (Shackel, 1984; Verbeek & Slob, 2006; Wever, Van Kuijk & Boks, 2008). Interaction problems, CPFs and the consumer's response to a perceived failure are dependent upon the characteristics of consumers using these products and upon the variability of the environments in which these products are used.

The combination of the insights from these theoretical models leads to the design of a conceptual research model to position the concept of CPF and its antecedents in the propagation of faults to consumer complaints. This model is shown in Figure 2.5. As will be further explained at the end of this section, this dissertation will limit its investigation to the highlighted relations within the "initial research focus".

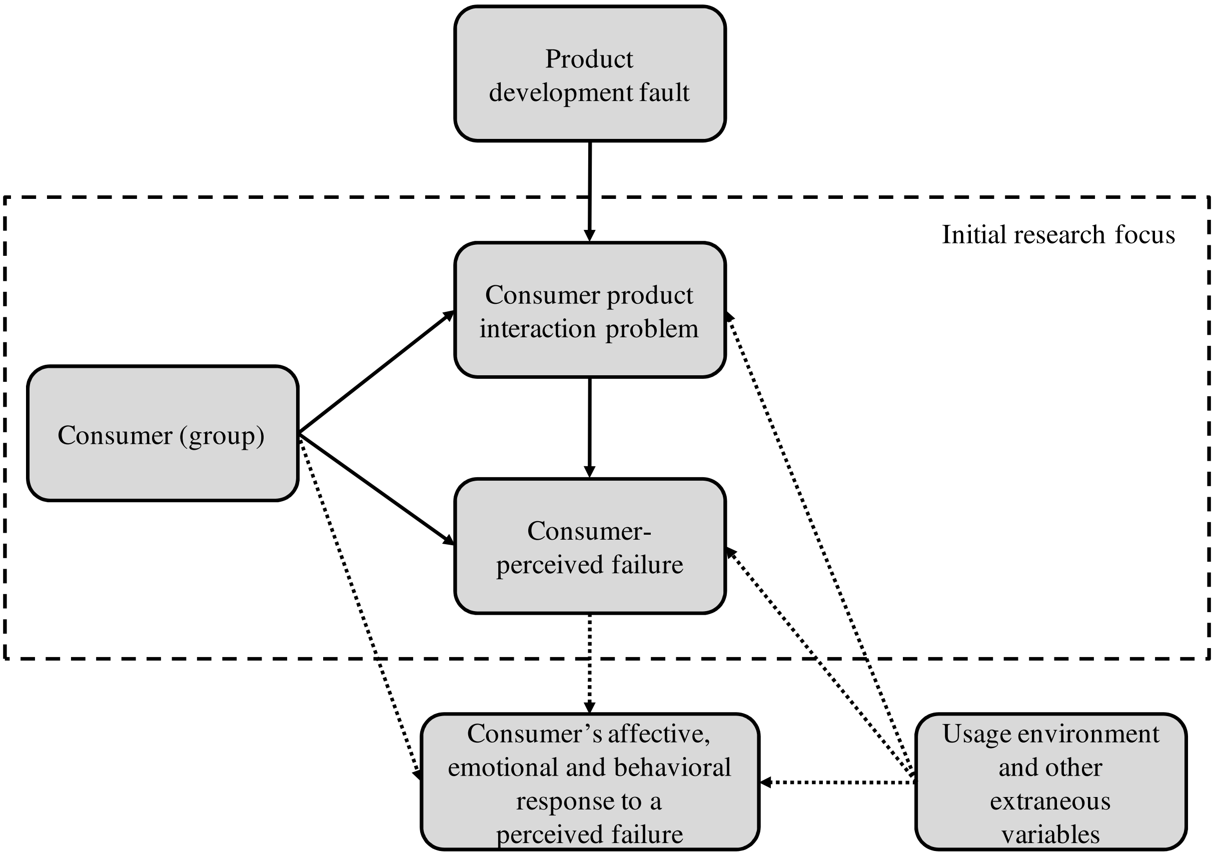


Figure 2.5 Fault-complaint propagation model

The functioning of this model can be explained as follows. Faults can, for certain consumers under certain usage conditions, lead to a problem when a consumer is interacting with the product. Subsequently, again depending on the consumer and the usage situation, psychological processing of this problem could result in a situation where the consumer perceives that something is actively wrong with the product, which s/he may then decide to report to the manufacturer or other parties involved in this usage situation (e.g. a service provider): *a consumer-perceived failure*. Next, again depending on the consumer, the usage situation and other external factors, psychological processing of this perceived failure results in an affective, emotional and behavioral response (e.g. a complaint).

Regarding the use and interpretation of the concepts and relations in this model, please note that, similar to the relations described in the Q&R fault propagation model, the top-to-bottom propagation of faults to the consumer's response to a perceived failure should be interpreted as such that the occurrence of a cause can possible, but not necessarily lead to the occurrence of its effect. For example, for some consumers with a certain usage profile in certain usage environments, a fault can lead to an interaction problem while for other consumers this problem never occurs (although the preceding fault is assumed to be always present). Additionally, in some instances the combination of several faults can lead to a problem during interaction (e.g. a software fault combined with feedback messages in the UI that are misinterpreted by the consumer). Finally, it is important to note that, although this dissertation focuses on faults as an antecedent, the propagation can start at (a combination of) any of the three described antecedents of CPFs.

As was described in Chapter 1, the goal of this dissertation is to investigate the relation between the heterogeneity of consumer groups using CE and the propagation of product development faults to consumer complaints. Due to time constraints and due to the context of the TRADER project in which this research takes place, only several aspects of this relation will be further investigated. First of all, the influence of usage conditions and other extraneous variables on the propagation is out of the scope of this research project and will not be further investigated. Secondly, since the goal of the TRADER project is to focus on product failures, this dissertation will specifically investigate the propagation of (software) faults to CPFs. In this context, input from the TRADER project will be used to identify and use relevant product development faults as a potential source of CPFs but the detection, analysis and prevention of these problems in the product (software) itself is dealt with in other TRADER projects (see also section 1.1.2 and Stroucken et al., (2005)).

2.2 Different views on consumer diversity

As stated in section 1.3, the aim of this dissertation is to gain more insight into the relation between the variability of consumers and the propagation of faults to CPFs. Following the definition of the antecedents and consequences of CPFs, the next step is to investigate which differences between consumers affect this propagation and if so, in which manner they affect this propagation. As a first step, in this section it is investigated how consumer diversity can be addressed to gain more insight into interaction problems and CPFs as described in the conceptual propagation model. In literature, many different views on how to address consumer diversity can be found. This section addresses these different views by first investigating which method is most appropriate to model consumer diversity for the goal of this research. Subsequently, based on the chosen method, different consumer models and underlying characteristics from both marketing and HCI perspective are discussed resulting in a choice to further focus on cognitive models.

In section 2.2.1, different methods to model consumer diversity are discussed. Subsequently, in section 2.2.2 it is discussed why differentiating on the target consumer does not fully cover differences in interaction problems and CPFs. Finally, in section 2.2.3, an overview is given of relevant consumer characteristics for HCI and design research and it is discussed why cognitive consumer models are the most important models for better understanding the variability in interaction problems and CPFs.

2.2.1 Different methods to model consumer diversity

Consumer diversity is addressed by many different models in marketing, HCI, design and consumer behavior research. Based on the research by Muller, Millen and Strohecker (2001), the following methods to model consumer diversity can be defined:

- The statistical average consumer: One statistical average consumer with certain characteristics stands for the whole consumer population for a certain product. Such an approach only works for very homogeneous and small consumer populations.
- Statistical stratified sample: Characterizations of consumers in a small number of relevant attributes and subsequently selecting representative consumers based on these attributes. This approach is suitable for heterogeneous populations but its quality depends on the assumptions underlying the selection and measurement of attributes.
- Strategic sampling for diversity: Continuous sampling of consumers groups until the most important sources of heterogeneity are exhausted. This approach is most suitable for discovering diversity but not for using these different consumer groups for further research.
- Politically representative consumers: Commonly used in participatory design and co-design approaches in which several consumers represent the interest of larger consumer groups during product design (see for example Grudin and Pruitt (2002) and Battarbee (2004)). These approaches are advocated for product design but are less suitable for large, heterogeneous populations with high levels of uncertainty.
- Fictitious consumers: Approach in which no real consumers are involved but instead in-depth descriptions of fictitious consumers who represent the target population are used. For example, in design research "personas" are used (Battarbee, 2004; A. Cooper, 1999; Pruitt & Grudin, 2003). A persona is defined as "a precise description of the user and what s/he wishes to accomplish" (A. Cooper, 1999, pg. 123). It is a hypothetical archetype of an actual user, described in terms of specific goals, operating in specific environments, having specific characteristics and skills. Personas should prevent product developers from designing a compromise that incorporates some useful aspect for every user but is not satisfying for any of them (A. Cooper, 1999). However, these aspects do not help to investigate how variability of consumers can affect interaction problems and CPFs because for such analysis it is important not to generalize into several hypothetical archetypes but to reflect actual usage and failure perception of dynamic consumer groups. In other words, personas and other

fictitious approaches are not capable of handling uncertainty with regard to consumer diversity.

- Extreme consumers: The use of “untypical” consumers to challenge the product design and develop new insights. For this approach two examples can be found. First of all, Von Hippel (1986) proposes to use so-called “lead users”. Lead users are users who face needs months or years before the mass consumer market encounters them and expect to significantly benefit by obtaining a solution to those needs (Herstatt & Von Hippel, 1992). This research is driven by the observation that insights of average consumers into new product solutions are constrained by their own real world experience and concept development and test methods are therefore unlikely to generate novel concepts that conflict with the familiar concepts (Von Hippel, 1986). Although this method is very useful for concept generation, it does not provide further insight into the variability of problems consumers encounter during product usage since lead users only represent a very small fraction of mass consumer markets. Furthermore, identifying lead users is very difficult (Kaulio, 1998). Secondly, besides lead users, in product development literature the use of extreme consumers is proposed in the High Contrast Consumer Test (HCCT). This test is designed to make consumer testing more effective by maximizing variability in the interaction between the product and consumers to provoke product failures early in the PDP, which would normally only be identified after usage in the field (Boersma, Loke, Loh, Lu & Brombacher, 2003). The HCCT includes use of extreme users who are on the edge of the defined target user profiles (Baskoro, Rouvroye, Brombacher & Redford, 2003). However, the literature on HCCT does not specifically address why certain demographics were chosen as discriminating factors between extreme user groups and does therefore not provide further insight.

Based on this overview, it can be concluded that statistical stratified samples are the most suitable to gain more insight into consumer diversity for the purpose of this research. Consequently, the emphasis of this research project is to select one or several important characteristics of consumer groups to gain a deeper insight into how this characteristic affects differences in interaction problems and CPFs. However, it is important to note that the goal of this research is not to draw general inferences for a general population for which statistical stratified samples are most often used in survey research.

Consumer characteristics can be differentiated from both a marketing and a HCI or consumer behavior perspective. The following section discusses the most commonly used differentiation of consumers from a marketing perspective: technology adoption.

2.2.2 Differentiating consumers based on technology adoption

In HCI and product design and development literature, the most commonly used sampling of consumers for product design and test methods is the selection of “target customers” or “target users” (Bekker & Long, 2000; Griffin & Hauser, 1993; Ozer, 1999). In other words, a product should be designed and tested with people who will probably be buying and/or using the future product. One of the most commonly used models is the model of innovation diffusion developed by Rogers (2003). Innovation diffusion is defined by Rogers (2003) as: "The process by which an innovation is communicated through certain channels over time among the members of a social system". He further states that there are seldom innovations that represent a superior alternative to the previous product that it replaces. Consequently, an innovation creates uncertainty in the minds of potential adopters about its expected consequences as well as representing an opportunity for reduced uncertainty in another sense, i.e. solving an individual's need or perceived problem. Among other things, this process influences the degree to which an individual is relatively earlier in adopting innovations. The time element of the diffusion process allows the generation of diffusion curves and subsequently the classification of adopters into categories. Rogers (2003) identifies five adopter categories, which are plotted on a bell-shaped innovation adoption curve (see also Figure 1.2) and can be defined as follows:

- Innovators: This group is the first to adopt a new innovation. Among other characteristics, they are very eager to try new ideas, have a substantial amount of financial resources and have the ability to understand and apply complex technical knowledge.
- Early adopters: This group is more substantial than the innovators and in most social systems has the greatest degree of opinion leadership. The role of the early adopter is to decrease uncertainty about a new idea by adopting it and then conveying the subjective evaluation to other peers in their social network.
- Early majority: This group adopts new ideas just before the average member of a social system. They follow with deliberate willingness in adopting innovations, but seldom lead like innovators.
- Late majority: This group adopts new products just after the average member of a social system. Innovations are usually approached skeptically by this group and they do not generally adopt innovations until most others in their social system have done so.
- Laggards: This group is the last in a social system to adopt an innovation. They tend to be suspicious of innovations and adopter categories like innovators. They tend to be very slow in the innovation decision process and also have limited resources to adopt innovations.

Although this model and other demographic and lifestyle based adoption models are widely applied in research and practice, characteristics of segments determined by marketing departments are often too narrow to encompass the diversity of potential consumers in relation to consumer behavior (Berkman & Erbuğ, 2005). Furthermore, as stated in section 1.2, due to the changing business context of the CE industry, it becomes far less feasible to define homogeneous target consumer groups with a specific use profile with a high level of certainty (De Marez & Verleye, 2004; Grudin, 1991; Kujala & Kauppinen, 2006). In other words, consumer profiling based on adoption related demographics only works if companies have to deal with consumers who know exactly what they need, if those needs can be coupled with stable consumer profiles, and if the product design sufficiently meets that profile. In the highly uncertain market for CE none of these requirements are met.

Now the question remains which consumer profiles give more insight into differences in interaction problems and CPFs for complex CE products. To answer this question, literature shows that regarding the selection of consumers for product design and consumer tests, several aspects need to be taken into account. First of all, several authors argue that the selection of consumers and the method by which consumers should be differentiated depends on the goal of the test (Gould & Lewis, 1985; Muller et al., 2001; Vredenburg, Isensee & Righi, 2002). For example, when trying to predict product performance in the market, a general population should be used, while for identifying problems with novice users only novice users should be invited for the test (Vredenburg et al., 2002). Secondly, research has shown that the differentiation of consumers on deeper level characteristics could improve the predictive power of product test methods (Dillon and Watson, 1996; Kujala & Kauppinen, 2006). Differentiation of consumers on characteristics is especially important for research on consumer products because these products are used by larger and more diverse populations than products for which traditional ergonomics and HCI research is performed (Berkman & Erbuğ, 2005). However, in this context it is important on the one hand not to underestimate consumer diversity since consumers have a different understanding of product functioning than designers, but on the other hand not to overestimate consumer diversity because taking into account all potential characteristics can distract the designer from important issues (Berkman & Erbuğ, 2005; Gould & Lewis, 1985; Nielsen, 1993; Potosnak, Hayes, Rosson, Schneider & Whiteside, 1986). Consequently, to potentially gain more insight into how the variability of consumers affects the propagation of faults to interaction problems and CPFs, differentiation of consumers on deeper level consumer characteristics needs to be investigated. How and which characteristics will be further investigated in this dissertation will be discussed in the following section.

2.2.3 Relevant consumer characteristics for further research

Since this research project closely relates to HCI and consumer behavior research, an overview of relevant consumer characteristics and consumer models used in these research areas is shown in Table 2.1 on the next page. Although the categories do not fully overlap and have different purposes, in general the characteristics discussed in these papers relate to the model of product functioning by Kanis (1998), who argues that consumer activities are a consequence of the consumer's perception, cognition and subsequent use actions. Consumer characteristics can be divided into perceptual, cognitive and physical characteristics respectively, complemented by psychographic and demographic models which capture the cultural, habitual and emotional differences between consumers (Hasdoğan, 1996).

Table 2.1 Overview of consumer characteristics and consumer models discussed in HCI and design literature

Literature	**Consumer characteristics and/or models**
Dillon and Watson (1996)	Cognitive science: models used to predict individual differences in information processing. Personality and cognitive style: models used to predict individual differences in personality (i.e. "traits or stable tendencies to respond to certain classes of stimuli in or situation in predictable ways") and cognitive style (i.e. "stable patterns of information processing that are displayed by an individual"). Psychomotor differences and skill acquisition: models used to predict individual differences in psychomotor performance and skill acquisition.
Hasdoğan (1996)	Physical models: models that represent mechanical and dimensional characteristics of the human body. Cognitive models: models that represent the human being's sensory and cerebral processing system, his characteristics and limitations relates to the elements of that system and the outcome of such processes (e.g. mental and sensory models). Consequence models: models that represent undesired outcomes from human-machine interaction such as accidents, errors etc. Psychosocial models: models that represent the emotional, cultural and habitual characteristics of humans (e.g. psychographic and demographic models).
Kanis (1998)	Sensory characteristics: human characteristics used for noticing a product's functionality or functioning (i.e. perception). Mental characteristics: human characteristics used for understanding a product's functionalities or functioning (i.e. cognition). Physical characteristics: human characteristics related to the ability to perform use actions when using a product (e.g. exerting force).
Kujala and Kauppinen (2004)	Personal characteristics: demographics, lifestyle, personality, emotions, attitudes, skills and physical abilities and constraints. Task related characteristics: goals and motivation, tasks, usage, training and experience. Geographic and social characteristics: location, culture and social connections, societies and organizations.

To determine which characteristics are of importance for this research, those characteristics need to be investigated that relate most closely relate to the variables of the fault-complaint propagation model. In the context of the increasing complexity of CE as discussed in section 1.1, designers rely more on the experience and ability of consumers, which are very diverse for large consumer populations (Berkman & Erbuğ, 2005). This specifically relates to cognitive functions such as problem solving, judgment, decision making and information processing in general (Kujala & Mäntylä, 2000; Roth, Patterson & Mumaw, 2002), which are central to understanding a product's functioning and subsequently interpreting potential problems and failures. Therefore, these characteristics are of most importance to investigate diversity in interaction problems and CPFs for large consumer populations. Nevertheless, other consumer characteristics such as demographics and potentially moderating or mediating factors will need to be taken into account when drawing conclusions.

2.3 Addressing consumer diversity by using cognitive models

To further understand the potential contribution of a differentiation on cognitive functions, in this section relevant research on the cognitive processing of a product's functioning is discussed. This section concludes with the presentation of an initial research model.

2.3.1 Usability and related research

In usability and related research fields, a substantial amount of research can be found on the effect of the consumer's cognitive characteristics and models on the consumer's understanding of a system and subsequent performance when using that system. In the ISO 9241-11 standard (ISO 9241-11, 1998), usability is defined as: "The extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use". These usability attributes are measured through various techniques of which usability testing with users is an important aspect. Depending on the goal of the test, users need to test the system for achieving specific goals in context dependent environments (Nielsen 1993, p. 27). In this context, usability testing is mainly focused on identifying interaction problems and subsequently assessing how severe these problems are. Although usability literature commonly advocates the differentiation on "typical" or "target" users as a main criterion for selecting users for usability tests (Battarbee, 2004; Ketola, 2002; Kujala & Mäntylä, 2000; Nielsen, 1993), according to Nielsen (1993, p. 28) another dominant way to categorize users for usability testing is "user expertise" which predicts learning. Three different dimensions of expertise are defined:

- Experience with the system
- Experience with computers in general
- Experience with the task domain

These dimensions are used to differentiate between novice, casual and expert users. A novice user has no or only minimal experience, a casual user is a person who uses the system intermittently and an expert user is a person who uses the system frequently. Although such a differentiation could give more insight into differences in interaction problems and CPFs, Dillon and Watson (1996) state that this basic differentiation lacks predictive power which could be improved by learning from research on individual differences in psychology.

In HCI literature, research can be found that relates this novice-expert differentiation to differences in the way in which consumers deal with errors encountered during product use (mostly related to computerized tasks). For example, research shows that deeper knowledge of how novice users perceive errors that occur during web browsing can be helpful for designers to lessen the occurrence or even the perception of occurrence of those errors (Lazar, Meiselwitz & Norcio, 2004; Lazar & Norcio, 2003). Especially because novice users lack expertise of the system, they make more errors during product usage, have more difficulty to recover from those errors and also could perceive these errors differently than expert users (Lazar & Norcio, 2003).

However, another study showed that differences in cognitive abilities not always predict the number of errors made during computerized tasks but are reflected in the types of errors encountered and in the way via which errors are resolved (Prümper, Zapf, Brodbeck & Frese, 1992). Having more experience with or expertise of a system does therefore not always result in fewer errors made.

Finally, it is interesting to look at the results of a study on differences between complaint scripts of experts and novices (measured through prior knowledge on the complaint process) conducted by Martin (1991). The results of this study showed that:

- Experts have significantly more important information, a better organized hierarchical structure and a higher level of abstraction in their complaint descriptions.
- Experts can abstract better to new complaints situations.

Consequently, the results of the studies above indicate that differences in cognitive abilities can result in differences in performance, error perception and subsequent response to those errors. Although research in this domain is mainly focused on computerized tasks, these concepts can also be interesting in the context of the fault-complaint propagation model.

2.3.2 Mental models

In HCI literature, a substantial amount of research related to cognitive science and cognitive models can be found which focuses on the concept of "mental models" in terms of its antecedents and consequences in human behavior, its measurements and its implications for product design. According to Van der Veer and Del Carmen (2002), "the interest in mental models from HCI is based on the idea that, by exploring what users can understand and how they reason about the systems, it is possible to design systems that support the acquisition of the appropriate mental model and to avoid errors while performing with them". In HCI literature a mental model is commonly known as "the mental representation constructed through interaction with the target system and constantly modified throughout this interaction" (Norman (1983) as cited by Van der Veer & Del Carmen, (2002)). Research shows that when consumers have such a model of a system it facilitates the consumer to exactly infer how the product works (Kieras & Bovair, 1984). For example, Uther & Hailey (2006) show that training the correct mental model of web browsing positively affects web browsing navigation performance. Consequently, one can also argue that by understanding a consumer's mental model of a product one can also better understand how and why interaction problems and CPFs occur.

However, mental models are not simply observable which makes them difficult to elicit, measure and subsequently differentiate consumers on these models (Zhang & Cignell, 2001). A mental model is not a characteristic of a consumer but an instantiation of consumer knowledge of a system (Van der Veer & Del Carmen, 2002) and is therefore not an easy segmentation tool. Furthermore, research has shown that mental models are considered to be incomplete, inaccurate and unstable (Norman, 1983; Staggers & Norcio, 1993; Thatcher and Greyling, 1998). Norman (1983, p. 8) adds that mental models do not have firm boundaries (people confuse different products) and are "unscientific" (people maintain superstitious behavior patterns to save physical and mental effort). In HCI literature, the concept of mental models is most widely applied in research on interaction with computers, although, due to the increasing complexity of the CE as described in section 1.1, the concept is equally relevant for consumer products (Van der Veer & Del Carmen, 2002).

Although mental models cannot be used as a consumer segmentation variable, research has shown that the structure and usage of these mental models depend on certain consumer characteristics, for example:

- Zhang and Cignell (2005) show that there is a significant effect of educational and professional status, academic discipline and computer experience on the mental model of information retrieval systems.
- Thatcher and Greylin's (1998) findings suggest that mental model categories may be hierarchically ordered (in order of level of detail and completeness) according to the consumer's experience with using the Internet.
- Ziefle and Bay (2004) show that age significantly affects the correctness of the mental model of a cellular phone menu.
- Docampo Rama (2001) argues that besides age differences, the so-called "technology generation" in which a consumer grew up (e.g. the electro-mechanical generation encompasses consumers born before 1960) also affects how consumers deal with errors in new systems.

Overall, research shows that the consumer's level of knowledge or experience with a device seems to be a strong determinant of a consumer's mental model of a system (Thatcher & Greyling, 1998; Van der Veer & Del Carmen, 2002) and subsequently affects problem solving and learning when interacting with complex systems (Staggers & Norcio, 1993). Moreover, Van der Veer and Del Carmen (2002) discuss that the consumer's knowledge and mental model of a system often do not match the knowledge needed to handle meaningfully in a certain situation. Consequently, this concept could be used to differentiate consumers and to gain better insight into the relation between consumers and the propagation of product development faults to CPFs in the context of complex CE.

2.3.3 Conclusion

In section 2.2.2 it was discussed that for a better understanding of the effect of consumer diversity on interaction problems and CPFs, cognitive consumer models are of the most relevance because they relate to the consumer's understanding of a system's functioning. In this section, it was shown that in HCI and consumer behavior literature differences in the consumer's knowledge of or experience with a complex system are commonly used to predict and analyze differences in understanding product errors and mental models of those systems. Although most of this research is conducted for computer systems, research by Arning and Ziefle (2009) has shown that these differences in cognitive abilities can also help to predict menu navigation performance of PDAs.

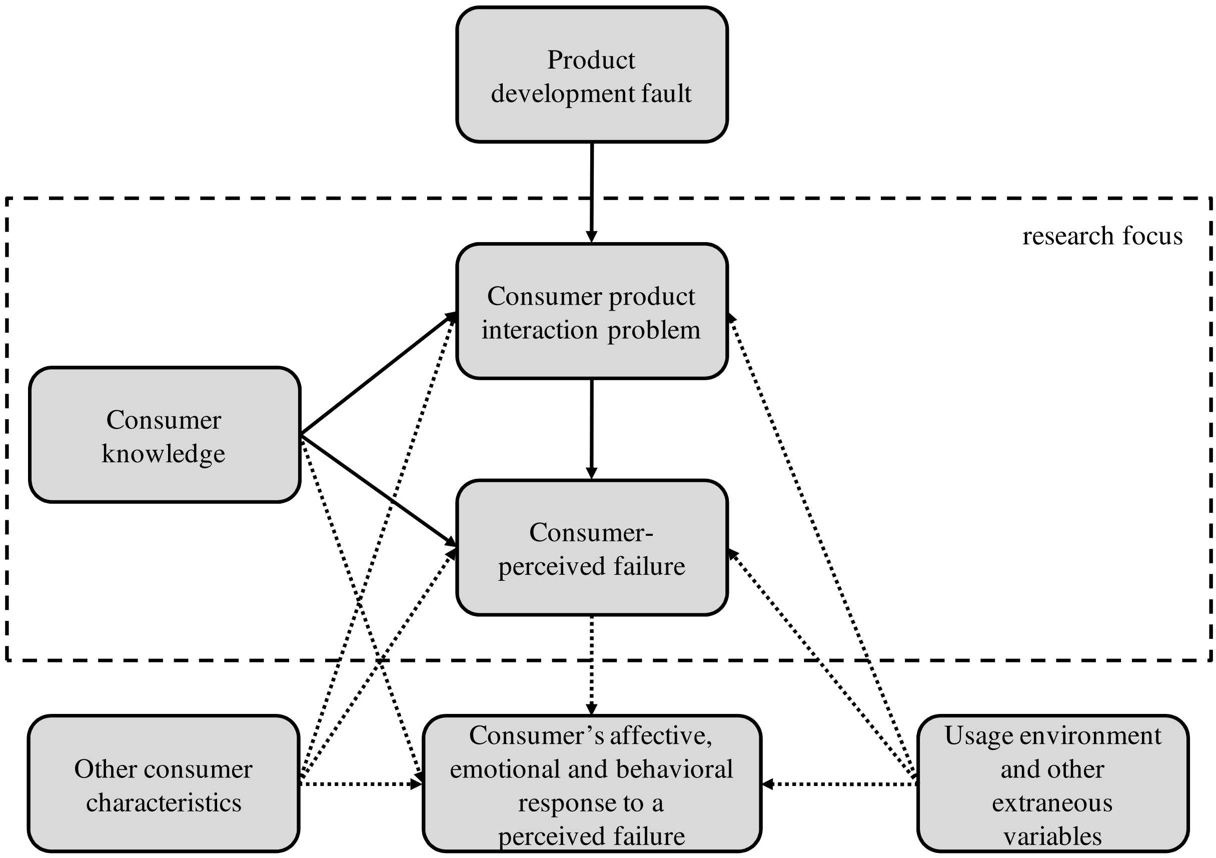


Figure 2.6 Initial research model

Although there is a substantial amount of research on consumer knowledge and its consequences for consumer behavior, most of this research focuses either on computerized tasks performance or on pre-purchase consumer behavior (for example, see Cordell (1997)). In this context it is interesting to investigate how differentiating on consumer knowledge can help to better understand the propagation of product development faults to CPFs for large and very diverse consumer groups in realistic product failure situations. Consequently, an explicit choice is made to further focus this research on the investigation of the effect of consumer knowledge on the occurrence of interaction problems and CPFs in complex CE. Nevertheless, as discussed in section 2.2.2, other relevant consumer characteristics such as demographics and other potential moderating and mediating variables need to be taken into account when evaluating the effect for large consumer groups. The initial research model is shown in Figure 2.6.

In Chapter 3, the research variables that are in the main research focus will be further elaborated upon followed by the research questions and the research approach of this dissertation.

3 Research Model

In the previous chapter it was discussed that this dissertation will investigate how consumer knowledge affects the propagation of product development faults to CPFs. However, the propagation model and its research variables are formulated at a high level of abstraction. This chapter further elaborates upon the selected research variables, its measurements and related relevant research and concludes with the research questions and research approach.

As a starting point, in section 3.1 the effect of an easy-to-apply measurement of consumer knowledge, familiarity, on the consumer's perception of a deliberately implemented fault in the teletext function of a TV is investigated. Based upon the results of this experiment, a more detailed literature review on the constructs and measurements of both consumer knowledge and CPFs is conducted. The results of this literature review are presented in section 3.2 and section 3.3 respectively. Subsequently, section 3.4 presents a detailed research model and the research questions addressed in the remainder of this dissertation. Finally, in section 3.5 the overall research approach for this dissertation is discussed.

3.1 Exploring the effect of familiarity on CPFs: Teletext experiment

3.1.1 Introduction

The propagation model and its research variables need to be further elaborated upon before a detailed research model and research questions are developed. As discussed by Stangor (1998, chapter 2), observation of behavior of people in the real world can be used to develop a theoretical model before defining research questions or hypotheses. For the purpose of this research project, a small-scale explorative experiment is therefore conducted as a starting point to gain more insight into the effect of consumer knowledge on the propagation of product development faults to CPFs and the underlying factors of this propagation. Because consumer knowledge consists of multiple components (Alba & Hutchinson, 1987; Cordell, 1997), in this experiment the most easy-to-use and most commonly used measurement of consumer knowledge, i.e. product familiarity, is used[6]. In this context, familiarity can be defined as (Alba & Hutchinson, 1987): "the number of product-related experiences that have been accumulated by the consumer".

In the experiment the effect of familiarity with teletext on the consumer's perception of a deliberately implemented failure in the teletext functionality of a Cathode Ray Tube (CRT)

[6] This construct of consumer knowledge is most commonly used in HCI literature to denote expert-novice differences (see also section 2.3.1).

TV is investigated[7]. Teletext is a text-based information service which is displayed on a TV by using the teletext button. Simply put, this information is broadcasted as digitally encoded data added to the cable signal (Limann & Pelka, 1991, p. 517). The content of this information is different for each broadcaster and/or channel. The main reasons for selecting this function were: 1) This was a relevant and realistic failure from a product development perspective, 2) it was relatively easy to implement a reproducible fault in this function, and 3) teletext is a common function in TVs which ensured consumers could be differentiated on usage experience.

Summarizing, this experiment is used to answer the following questions:

- How does the consumer's teletext familiarity affect the consumer's perception of the cause of a product failure in that function?
- How does the consumer's teletext familiarity affect the consumer's workaround strategy after experiencing a failure in that function?

In the following section, the method used to answer these questions is discussed.

3.1.2 Method

To answer the questions defined above, an explorative between-subjects experiment is set up in which the effect of the consumer's teletext familiarity on the consumer's response to an implemented failure in the teletext functionality is observed.

Experimental variables

The independent variable under study is the level of familiarity with the teletext functionality. For this experiment, the participants were divided into two groups based on their level of usage experience of the teletext functionality. Teletext usage experience (as suggested by research on familiarity measurements by Smith, Caputi, Crittenden, Jayasuriya & Rawstorne (1999) and Söderlund (2002)) was measured by the frequency with which the participants accessed different kinds of information via teletext and the frequency with which the participants use teletext on different channels. The questions (in Dutch) can be found in Appendix 3.1.

As dependent variables effectiveness, perceived failure cause and applied workaround strategy were recorded. Additionally, the participant's mental model of the technical functioning of the teletext function was investigated. Due to the exploratory nature of this experiment, the perceived technical functioning of teletext and the perceived failure cause were addressed by open questions. The effectiveness was assessed based on the ability to complete the task with the implemented failure and the applied workaround strategy was measured through observation.

[7] Please note that this experiment is the same explorative experiment as described by De Visser (2008, chapter 4), but for this dissertation other measurements and analyses are used.

Apparatus and materials
A project team of ten TV system experts (both product developers and testers) selected and designed a failure scenario in the teletext functionality of a CRT TV. According to these experts this is a realistic failure scenario that is caused by (the software of) the TV. The scenario involves that after accessing one of the teletext pages, the page appeared to be (partially) black and did not give the required information. An example of a correctly displayed teletext page and the same page as it appeared in the failure scenario is shown in Figure 3.1.

In order to "solve" this failure, i.e. make the failure disappear, three different solvability levels were implemented in the TV that could be selected by the experimenters before an experiment. These solvability levels were:

- Switch to TV mode and back to teletext.
- Switch TV channel.
- Go to standby and back to TV mode (or switch the TV off).

The participant was not made aware of these solvability levels, but they were implemented to build-in real life failure scenarios. During normal product usage, the occurrence of a failure would probably trigger work around strategies by the consumer and depending on the cause of the failure these strategies are assumed to be different.

Figure 3.1 Example of normally functioning teletext (left) and the same teletext page in the failure scenario (right) (De Visser, 2008, p. 52, snapshot from NOS teletekst (2007)).

Sample
For the selection of suitable test participants, an electronic questionnaire was used. The participants for the experiment were selected based upon diversity in demographics, ownership and usage of a TV and teletext familiarity. An overview of the questions used in this questionnaire to differentiate participants on teletext usage experience is shown in Appendix 3.1. Based on available time and budget (in view of the explorative goal of this

experiment), the questionnaire was sent to 45 individuals of whom 35 filled it out. Subsequently, 29 respondents aged between 21 and 66 years volunteered to take part in the experiment. An overview of the respondent characteristics is shown in Table 3.1.

Table 3.1 Respondent characteristics in terms of age, educational level and gender

Age	*n*	%	Educational level	*n*	%	Gender	*n*	%
< 35 years	18	62.1	Lower	13	24.1	Male	15	51.7
35 years >	11	37.9	Higher	16	55.2	Female	14	48.3
Total	**29**	**100.0**		**29**	**100.0**		**29**	**100.0**

Respondents were recruited through friends, family, colleagues and students. This ensured a high response rate and willingness to participate in the experiment. The major drawback of this recruitment method is that some participants might be slightly biased due to their relation with the experimenters. This possible effect has to be taken into account when assessing the validity of the results of the experiment. The test participants were divided into a low and high teletext familiarity group based on a split on the mean score of the summated score on teletext usage experience. A separate pair-wise Mann-Whitney U test (with the level of significance set at $p=0.05$) (Mann & Whitney, 1947) confirmed that teletext usage experience is significantly different between the two groups ($p < 0.001$). In the low familiarity group 8 women and 7 men with a mean age of 36.9 years participated. In the high familiarity group 6 women and 8 men with a mean age of 31.4 years participated. Separate pair-wise Mann-Whitney U tests showed no significant differences between the participants' characteristics of the two groups in terms of age ($p \gg 0.05$), educational level ($p \gg 0.05$) and gender ($p \gg 0.05$).

Procedure

The experiment was performed in the consumer test facility of the research group on the university campus. This laboratory consists of two rooms, one which resembles a living room in which the test participant performed the test and the other in which test participants could be observed by the experimenters through a one-way mirror. Each experiment was recorded with two cameras, a non-obtrusive camera mounted on the ceiling to capture the consumer's actions and a camera on a tripod which was used to capture the consumer him/herself.

The experimental procedure consisted of the following steps:

1. Introduction to the experiment. To prevent potential bias the participants were explained that the test was used for improving the usability of future generation TVs without referring to the precise goal of the test and they were assured that natural behavior was the most important. Finally, the experimenter made sure that the participant fully understood the task list and explained that there were no time and task completion restrictions.

2. Short interview with the test participant to retrieve the participant's mental model of how they perceive teletext functions from a technical perspective.
3. Experiment with the teletext functionality. The test participant was asked to complete three different tasks that involved retrieving information via teletext. The three tasks can be summarized as follows (the complete task list can be found in Appendix 3.2):
   - Search for the latest results of the football matches in the Dutch football competition.
   - Search for a movie you would like to see this evening through the teletext program guide.
   - Look up the arrival time of flight number KL 1168 from Helsinki, which is due to arrive today at Schiphol airport in Amsterdam.

The first two tasks were used to get the test participant acquainted with the TV and its remote control, the test conditions and the tasks. During the final task the teletext failure scenario was triggered in which the participant was initially not able to retrieve the requested information due to the implemented fault. The flight information could only be found on teletext page number 758 which, on access, appeared black (besides the page number) as shown in Figure 3.1. Depending on the selected failure scenario, the participant could solve this problem by switching the teletext page, switching TV channels or switching the TV off/on. The participants of the low and high familiarity groups were distributed equally across these different solvability levels. During the experiment the participant was asked to think aloud which enabled the researchers to capture (a part of) the process the users goes through when performing the tasks (e.g. why s/he takes a certain step).

4. After the experiment, the participant was asked to fill out a questionnaire that was used to measure the User Perceived Failure Severity (UPFS), i.e. the level of user irritation caused by the failure (De Visser, 2008, chapter 4). The analyses of these measurement are discussed by De Visser (2008, chapter 4) and are beyond the scope of this dissertation. Subsequently, the participant was debriefed about the goal of the test and the purpose of the research project.

Each experiment took on average 15 minutes. Afterwards, each participant was rewarded with a gift voucher worth €15,-.

3.1.3 Results

In this section, the results of the experiment with regard to the effect of teletext familiarity on the consumer's perception of the failure cause and the applied workaround strategy after experiencing the failure scenario are discussed. In view of the exploratory nature of this experiment, mainly qualitative comparisons of the dependent variables are used. For the statistical tests the level of significance was set at $p = 0.05$. Results within the less restrictive level of $p = 0.1$ are indicated as marginally significant.

Mental model

The results of the coding of the participants' perception of the technical functioning of teletext shown in Figure 3.2 gave some interesting insights. First of all, no differences in mental model can be observed between low and high levels of familiarity with teletext. Secondly, there is a broad scope of answers that can be categorized in five different groups, as shown in Figure 3.2. Based on the categorization of mental models in this figure, it can be concluded that only three out of the 29 participants have a reasonably correct mental model of the functioning of teletext.

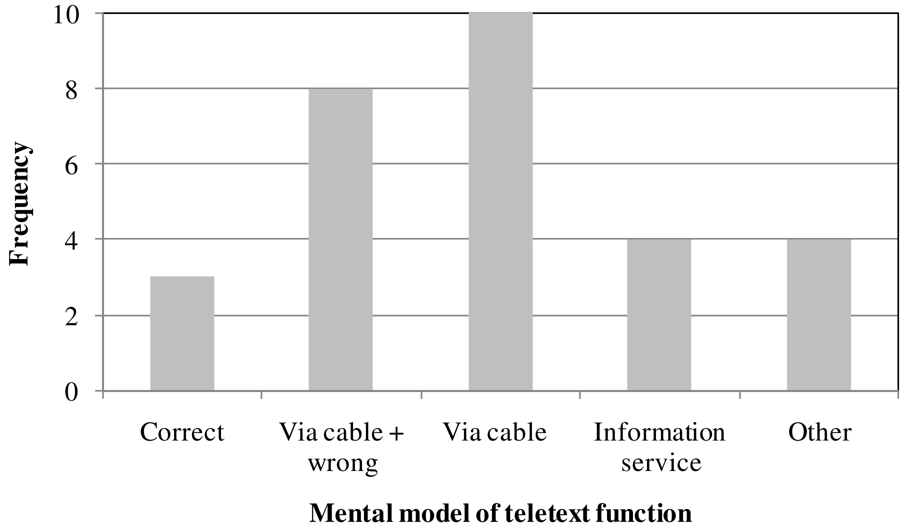


Figure 3.2 Overview of the mental models on the functioning of teletext

The different categories in this graph can be described as follows (in decreasing order of correctness):

- Correct: the participant perceives the technical functioning of teletext in accordance with the real technical functioning as described in section 3.1.1.
- Via cable + wrong interpretation of technical functioning: the participant understands that teletext is information sent to the TV via the cable together with the TV signal and the participants thinks s/he understands the underlying technical principle but is incorrect.
- Via cable: the participant understands that teletext is information sent to the TV via the cable together with the TV signal, but does not know the technical principles behind it[8].

[8] Please note that the difference of this category with the higher level "via cable + wrong..." category is that on this level the participant was not able, neither correct nor incorrect, to reason about how the information sent via the cable is translated into a teletext image on the TV screen.

- Information service: the participant understands that teletext is information sent by an external party, but does not know how this information is transferred to the TV.
- Other: the participant does not know how teletext functions and/or gives a completely incorrect answer (e.g. "teletext is sent to the TV via a direct link with a satellite").

Effectiveness

Overall, the results of a Mann-Whitney U test show that only a marginally significant effect of teletext familiarity on the ability to complete the task in the failure scenario can be observed ($p < 0.1$). In other words, there is an indication that consumers in the low familiarity group are less able to complete the failure scenario task (5 out of 15 completed the task) than consumers in the high familiarity group (10 out of 14 completed the task). Furthermore, the results of a Kruskal-Wallis test show significant differences for the effect of the solvability level on task completion (χ^2 (2) = 6.49, $p < 0.05$). However, separate pair-wise Mann-Whitney U tests (after applying Holm's sequential Bonferroni correction (Howell, 2002)) show no significant differences between the individual scenarios.

Consumer perception of the cause of the failure

The most striking result of this experiment is the number of participants who perceive the cause of the failure to something else than the TV (25 out of 29). In other words, the majority of the participants have a different perception of the failure cause than the TV system experts who chose and designed the failure scenario. An overview of the perceived causes and differences between the low and high familiarity groups is shown in Figure 3.3. From the results presented in this figure it can be seen that there are only slight differences between the low and high familiarity groups in terms of frequency with which the different perceived causes are mentioned.

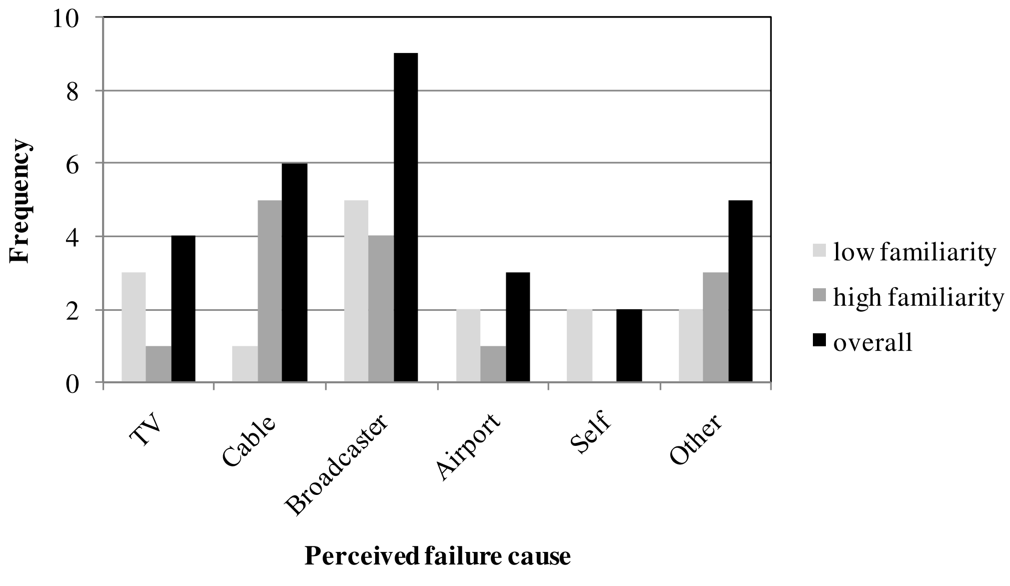


Figure 3.3 Overview of perceived failure causes

The results of Mann-Whitney U tests also show no significant differences between the low and high familiarity groups and no significant effect of the level of correctness of the mental model on the perceived failure cause.

Applied workaround strategy

Finally, the applied workaround strategy after encountering the failure during the final task is discussed. The results presented in Figure 3.4 show differences between low and high familiarity groups. Please note that in this figure the total sum of the workaround strategies is higher than the number of participants because some participants applied multiple strategies. Overall, the participants in the low familiarity groups remain on the same channel and try to overcome the problem by changing the teletext page and switching the teletext function off and on. In contrast to this strategy, the participants of the high familiarity group mainly switched channels to try to overcome the problem by accessing the same teletext page on a different channel. Statistically, the results of Mann Whitney U tests only show a significant difference for switching channels as a workaround strategy ($p < 0.01$). Please note that this effect is not due to the different solvability levels implemented in the failure scenarios because the participants were equally assigned to these different levels.

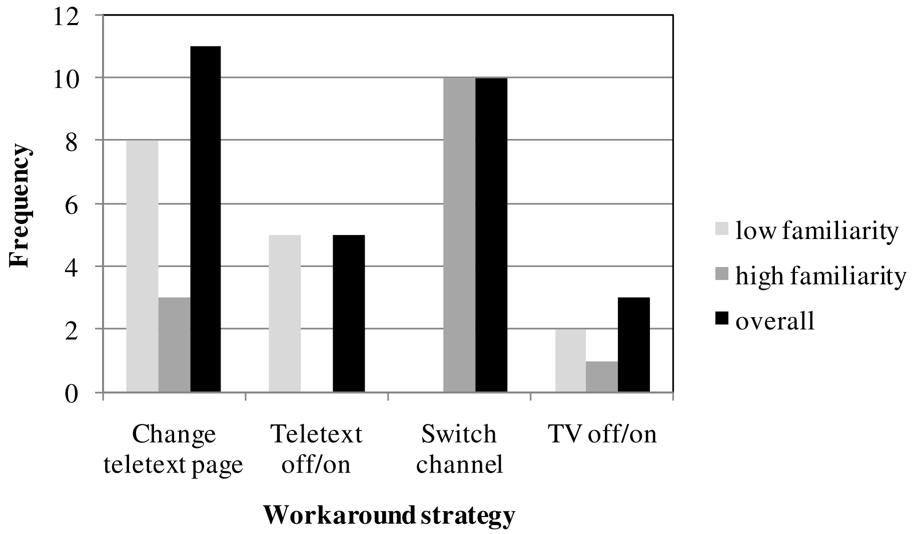


Figure 3.4 Overview of applied workaround strategy

3.1.4 Conclusions

From the results of this explorative experiment, more insight is gained into the (measurement of) potential effect of consumer knowledge on the propagation of product development faults to CPFs and its underlying factors. First of all, this experiment successfully demonstrated the ability to use simulated failure scenarios in a CRT TV used in a laboratory setting to evaluate the consumer's reaction to product failures. Although the laboratory setting allowed for control of external factors that prevent the ability to analyze the consumer's response to a failure in real-life settings reliably, one has to take into account that a laboratory setting could provoke a different response to a failure compared to a real-life setting.

Secondly, several conclusions can be drawn regarding the effect of familiarity on differences in the consumer's perception of the failure cause and the applied workaround strategy. The results show that there is no significant effect of teletext familiarity on the consumer's perception of the cause of the product failure. This conclusion is further strengthened by the observation that differences in teletext familiarity did not result in differences between the mental models of the teletext function.

However, the results do show a large diversity in both the perceived failure causes and perceived technical functioning of teletext. Furthermore, the results show an effect of teletext familiarity on the effectiveness in task completion and on the applied workaround strategy after encountering the product failure. Two explanations for these results can be found. First of all, the relatively simple measurements of both product usage and failure cause perception may not accurately reflect the underlying factors. As such the construct validity (i.e. whether the instruments to measure both product usage and failure perception are the best ones for measuring them) can be questioned (Goodwin, 2005, p. 116). Secondly, the differentiation of consumer knowledge by using product familiarity measurements may not represent differences in deeper levels of product understanding. Research shows that although product-related experience is expected to improve the ability to use a product, it is not a sufficient condition for product expertise, which relates more to the *understanding* of a product (Alba & Hutchinson, 1987; Cordell, 1997).

Consequently, the results of this experiment show that for a better understanding of the effect of consumer knowledge on the propagation of faults to CPFs, additional literature review is needed on the following aspects:

- The factors underlying "perception" of failures and its measurements.
- The factors underlying a consumer's response to perceived failures.
- The measurement of consumer knowledge differences.

Finally, several other limitations of this experiment could have influenced the results. Although differences between the two familiarity groups were to some limit controlled for (see sample discussion in section 3.1.2), the use of a convenience sample might have influenced the results leading to measurement errors. A larger random sample could reduce this effect. Furthermore, in hindsight one can question the validity of the selection of the failure scenario. From the results of the exit questionnaire it can be observed that more than 93% of the participants disagree with the statement that this TV would have to be brought back to the shop. Even more than 96% of the participants would not consult the helpdesk of the TV manufacturer. Additionally, more than 50% of the participants stated that, when asked what they would do to acquire the information described in the task list when they would be at home, is to search for it on the Internet. Consequently, the teletext function and the implemented failure as described in the failure scenario do not seem to be important from the perspective of the selected participants whereas the DTV system experts considered this failure to be relevant. A failure in a more important function from a consumer perspective might have resulted in a different perception of the failure cause and subsequent workaround strategy.

Nevertheless, the results of this experiment are a starting point to gain more insight into the effect of consumer knowledge on the propagation of product development faults to CPFs. Furthermore, the results indicate the need for additional literature research to further specify the research model, its variables and the research questions. In the following section, consumer knowledge constructs and measurements are discussed.

3.2 Consumer knowledge

As discussed in Chapter 1, in the case of complex technical products, consumers are less and less aware of and trained in the technical functioning of a product and therefore could have a not sufficient level of knowledge to understand a product's functioning. According to Engel, Blackwell and Miniard (1995), consumer knowledge can be defined as: "Information stored within memory". Especially in the research fields of Consumer Behavior, Psychology and Marketing, a whole body of research can be found on consumer knowledge and its antecedents and consequences.

This section will discuss some of the most important findings of this research area to better understand its hypothesized effect on usage behavior and CPFs. First of all, in section 3.2.1 the different consumer knowledge constructs are defined and discussed. In section 3.2.2, different measurements of these constructs are discussed. Finally, in section 3.2.3, conclusions are drawn on the use of consumer knowledge constructs and measurements in this research project.

3.2.1 Consumer knowledge constructs

In the field of consumer behavior research Alba and Hutchinson (1987) are the first to propose that consumer knowledge is a multidimensional construct consisting of two major components: *familiarity* and *expertise*. Familiarity is defined as (Alba & Hutchinson, 1987): "The number of product-related experiences that have been accumulated by the consumer". In this definition, product-related experiences should be regarded in a broad context, including advertising exposures, information search, interaction with salespersons, choice and decision making, purchasing and product usage in various situations. Next, expertise is defined as (Alba & Hutchinson, 1987): "The ability to perform product-related tasks successfully". When explaining this definition, Alba and Hutchinson (1987) further state that expertise should be regarded in a broad sense that includes both the cognitive structures (e.g. beliefs about product attributes) and cognitive processes (e.g. decision rules for acting on those beliefs) required to perform product-related tasks successfully. In other words, expertise relates to factual knowledge and familiarity relates to the level of contact with a product class (Cordell, 1997).

Consequently, an increase in product familiarity results in an increase in expertise. Alba and Hutchinson (1987) argue that there are five qualitatively distinct aspects of expertise that can be improved as product familiarity increases:

- Cognitive effort and automaticity: Simple repetition (which is an increase in familiarity) improves task performance by reducing the cognitive effort required to perform the task. In some cases repetition can lead to performance that is automatic.
- Cognitive structure: When familiarity increases, the cognitive structures used to differentiate products become more refined, more complete and increase the ability to represent products in terms of deep, rather than surface, structure.
- Analysis: The ability to analyze information, isolating that which is most important and task-relevant, improves as familiarity increases.
- Elaboration: The ability to elaborate on given information, generating accurate knowledge that goes beyond what is given, improves as familiarity increases.
- Memory: The ability to remember product information improves as familiarity increases.

From a consumer behavior research perspective, these five aspects of expertise are originally hypothesized to influence pre-purchase and purchase related consumer behavior constructs. However, it is hypothesized and to some extent also proven that the same constructs also influence product usage behavior (Alba & Hutchinson, 1987; Mitchell & Dacin, 1996; Shih & Venkatesh, 2004) and post purchase product evaluation (see for example research by Sujan (1985) on how expertise affects the product evaluation process).

It is important to consider that different tasks require different types of expertise and even for the successful performance of any particular task generally more than one type of knowledge is required (Alba & Hutchinson, 1987). For example, besides the distinction between expertise and familiarity, on a higher level a distinction can be made between two different categories of knowledge, primary base domain knowledge / core knowledge and supplementary base domain knowledge / supplemental knowledge (Moreau, Lehmann & Markman, 2001; Saaksjarvi, 2003). Moreau et al. (2001) have shown that, concerning the adoption of innovative products, the influence of consumer knowledge is different for continuous and discontinuous innovations. Almost every innovation can be derived from an existing product area. The existing product category is defined as the primary base domain (Moreau et al., 2001). Knowledge of this domain is used to learn about and develop a representation of a new product. When a new product is very similar to a previous product this process is straightforward. However, for discontinuous innovations (e.g. from film-based cameras to digital cameras) knowledge on the primary base domain does not necessarily increase understanding of the new product. For this type of innovation knowledge of additional (i.e. complementary) domains might influence the adoption process by filling in the "gaps" in knowledge caused by the difference between the new product and the related previous product generations (Moreau et al., 2001). For example, knowledge on computers and graphics software can help consumers to better understand digital cameras.

In conclusion, in this section several constructs relating to consumer knowledge were defined. An overview of these constructs and their relationships is shown in Figure 3.5. From this figure it can be seen that both expertise and familiarity are part of consumer knowledge. Regarding consumer knowledge, distinction can be made between core domain and supplemental domain knowledge. In the next section it will be discussed how these consumer knowledge constructs can be measured.

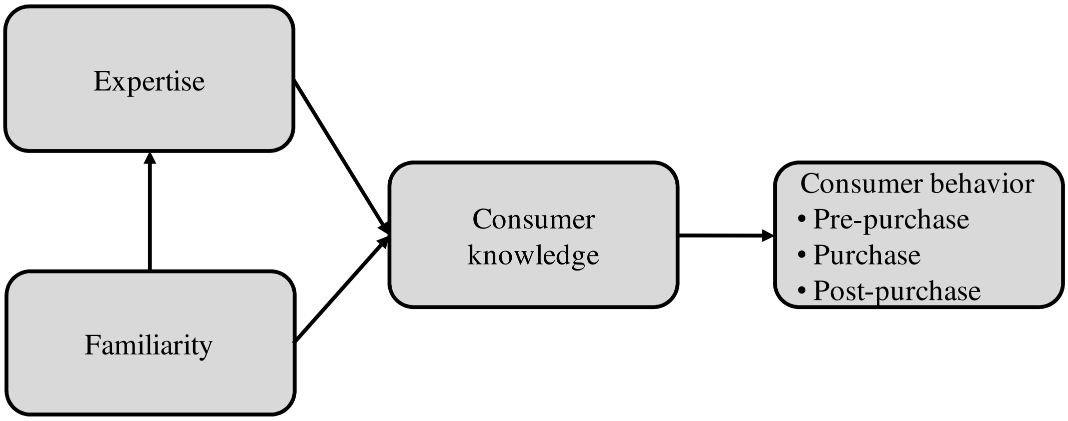


Figure 3.5 Overview of relations between consumer knowledge constructs

3.2.2 Measuring consumer knowledge

Traditionally, consumer knowledge has been treated as a one-dimensional construct (Alba & Hutchinson, 1987). Nowadays, in literature a distinction is made between objective and subjective measurements of both familiarity and expertise (Carlson, Vincent, Hardesty & Bearden, 2009). Although research has shown that these measurements are correlated, several authors also argue that different measurements affect different aspects of consumer behavior (Cordell, 1997; Dodd, Laverie, Wilcox & Duhan, 2005; Mitchell & Dacin, 1996). In the following each of the measurements and its applications are briefly discussed. Finally, conclusions are drawn for the further use of these different measurements in this research project.

Familiarity
In the past, familiarity and experience have often been used as proxy measures for consumer knowledge. Experience is often measured by self-report items (i.e. a subjective measurement) regarding the information search, ownership and usage of the product under study (Alba & Hutchinson, 1987; Dodd et al., 2005; Park, Mothersbaugh & Feick, 1994). Alternatively, experience can also be objectively measured by measuring the totality of the externally

observable direct or indirect interactions (e.g. amount of use, diversity of use and sources of information used) with a product across time (Smith, Caputi et al., 1999; Söderlund, 2002). Although research has shown that product familiarity contributes to the ability to perform product related tasks it does not encompass all relevant aspects of consumer knowledge (the five aspects discussed in the previous section). Brucks (1985) and Cordell (1997) argue that the experience-based component of knowledge is less directly linked to behavior than the measurements of expertise. Product knowledge can originate from many sources and is not necessarily correlated with experience (Johnson & Russo, 1984) which could be specifically relevant for complex products such as consumer durables (Raju, Lonial & Mangold, 1995). In other words, different people behave differently despite having the same experiences; experience can only influence behavior when this experience results in differences in memory.

Subjective expertise

Subjective expertise is defined as (Flynn & Goldsmith, 1999): “a consumer's *perception* of the amount of information they have stored in their memory”. Subjective expertise is usually easier to measure than objective expertise (Brucks, 1985). While the measurement of objective expertise requires an individual test for each product type, the measurement of subjective expertise can be done with a standardized scale. For example, Flynn and Goldsmith (1999) developed and validated a scale consisting of five items for measuring subjective expertise (on a seven-point Likert scale).

Brucks (1985) argues that subjective and objective expertise are distinct concepts. Subjective expertise includes an individual's degree of confidence in his/her knowledge while objective expertise refers only to what an individual actually knows. Alternatively, Arning & Ziefle (2009) state that subjective expertise ratings only reflect quantitative and not qualitative aspects of expertise. Nevertheless, Park and Lessig (1981) argue that subjective measurements of expertise may better define consumer strategies and heuristics because they are based upon what the consumer thinks s/he knows. Consequently, as discussed above for experience, subjective expertise may not be a valid measure of what is actually stored in memory as well but it can have a significantly different effect on consumer behavior than objective expertise.

Objective expertise

Finally, objective expertise is defined as (Brucks, 1985): “the actual amount of information stored in memory”. According to Brucks (1985), an objective expertise measurement should include the following aspects:

- Terminology: Does the consumer know the meaning of terms commonly used in consumer reports and manuals.
- Available attributes: Is the consumer able to list the attributes which are common for products in this product class.

- Criteria for evaluating attributes: Is the consumer able to list the criteria on which product quality can be evaluated and is s/he able to list dichotomous attribute criteria (i.e. the attribute is either present or absent) and continuous attribute criteria (i.e. multiple levels of the attribute are possible.
- Attribute covariation: Is the consumer able to explain the relationship between product attributes (e.g. price and size versus other attributes).
- Usage situations: Can the consumer explain how a situation affects the choice for a product.

Alternatively, Popovic (2000) states that knowledge and knowledge representations (related to objective expertise) for evaluating differences in product usability can consist of as many as nine different categories (e.g. declarative, procedural, interface etc.). Information sources for developing questions on these aspects could be discussions with product experts and consumer reports. The measurement of objective expertise has some important implications. Objective expertise is idiosyncratic with respect to a specific product class, i.e. a test for objective expertise on product class A cannot be used in a test for objective expertise on product class B (Cordell, 1997). In other words, objective expertise is regarded as the most reliable measure of what people actually know but is also the most time-consuming and most difficult measurement. Furthermore, Park et al. (1994) state that a limited number of items in an objective expertise measurement scale cannot accurately represent an entire product domain. The goal of the test must therefore be explicitly defined to be able to develop suitable objective expertise items.

3.2.3 Conclusion

In literature many different definitions and measurements can be found for consumer knowledge, even within the generally agreed upon categories of objective and subjective expertise. In general, objective expertise measurements are argued to be most closely related to the consumer's understanding of a product but research results so far are not consistent. Park et al. (1994) therefore argue that multiple knowledge constructs must be considered and measurement appropriate for the given research context and research questions must be selected. For example, Carlson et al. (2009) have shown that for some product domains and measurement scales of consumer knowledge, the more easy to use subjective expertise rating can be used as a reliable surrogate for objective expertise.

Furthermore, several authors argue that it is important to include knowledge calibration when investigating the effect of consumer knowledge (Alba & Hutchinson, 2000; Carlson et al., 2009; Pillai & Hofacker, 2007). In this context knowledge calibration refers to the relation between confidence and accuracy in knowledge rather than only accuracy (Alba & Hutchinson, 2000). For certain situations it is hypothesized that miscalibration of knowledge (i.e. a consumer scores significantly higher on subjective expertise than on objective expertise) can have a more significant effect on for example decision making than the level of knowledge itself (Alba & Hutchinson, 2000; Carlson et al., 2009).

Finally, research has shown that caution needs to be taken when using consumer knowledge as a predictor of consumer behavior. According to Alba and Hutchinson (1987), effects of knowledge on consumer behavior can only be regarded as main effects and must be studied with context dependent moderating variables.

3.3 Failure attribution

This section further investigates literature to better understand underlying constructs and measurements of CPFs. In section 3.3.1, the consumer behavior model of the psychological processing of performance outcomes (as briefly discussed in section 2.1.3) is further elaborated upon to identify relevant antecedents of CPFs. In this section it is discussed why *failure attribution* is an important measurement of the consumer's perception of product failure causes. Subsequently, in section 3.3.2 the construct, measurements, antecedents and consequences of failure attribution are further investigated. Finally, in section 3.3.3 conclusions are drawn on the use of failure attribution as an antecedent of CPFs in this research project.

3.3.1 Understanding psychological processing of interaction problems: the consumption processing model

As discussed in section 2.3, this dissertation focuses on gaining more insight into how consumer knowledge differences affect the propagation of faults to CPFs. To further understand the underlying process, the psychological processing of performance outcomes as modeled in the general consumer processing model defined by Oliver (1996) (see also Figure 2.4) needs to be further discussed. This model is selected because it combines several theoretical perspectives on antecedents and consequences of (dis)satisfaction and offers a complete overview of factors influencing a consumer's response to experienced product outcomes. An overview of this model is shown in Figure 3.6.

Basically, the model consists of two different phases of consumption processing: the non-processing phase of consumption and the processing sequence phase of consumption. The non-processing phase of consumption, visualized by the relation between outcomes, primary evaluation, primary affect and satisfaction, relates to the consumer's reaction to consumption outcomes with more or less spontaneous affect. In other words, this is a primary appraisal resulting from a general observation that the product outcome was "good for me" or "bad for me".

Of interest for this research project is the processing sequence of consumption because this process shows a cognitive perspective of the processing of product outcomes. As discussed in section 2.1.4, the affective, emotional and behavioral response to a perceived failure is outside the scope of this research project. Consequently, of specific interest in this model is the relation between product outcomes, expectation-disconfirmation and attribution.

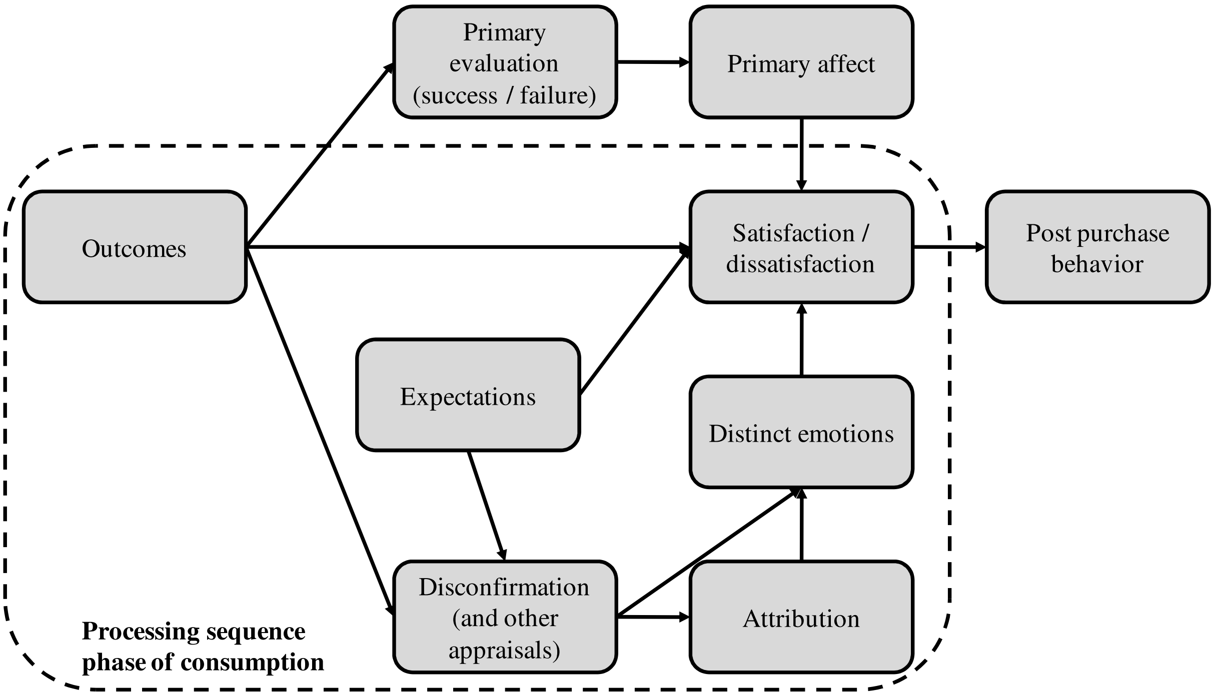


Figure 3.6 General consumption processing model (Oliver, 1996)

To show how these constructs can be used to gain more insight into CPFs, they first need to be defined as used in the general consumption processing model. Expectancy-disconfirmation refers to the discrepancy a consumer perceived when comparing an actual product performance with expectations, needs or other standards (Oliver, 1996). When consumers are subsequently primed to reflect on this discrepancy, as in why the consumption outcomes occurred in the manner they did, and generate reasons or assign responsibility for these outcomes, this process is referred to as attribution (Oliver, 1996).

In consumer behavior and marketing literature, attribution theory and research is well-founded. According to Folkes (1988), attribution research is concerned with all aspects of causal inferences, i.e. how people arrive at causal inferences, what sort of inferences they make, and what the consequences of these inferences are. Attribution theory is a combination of several theories that share core assumptions (Folkes, 1988; Silvera & Laufer, 2005): Heider's (1958) theory of naïve psychology, Kelley's (1967) covariation theory and Jones and Davis' (1965) correspondent inference theory. The limitations of these theories (Silvera & Laufer, 2005) were addressed by Weiner (1985, 1986) in a framework in which he proposed that besides the locus dimension (dispositional vs. situational or in other words internal vs. external), also controllability and stability are additional attributional dimensions. Up to now this framework has been one of the most frequently used models and has been influential in many different research areas (Oliver, 1996; Silvera & Laufer, 2005). Most of the applications of attribution research can be located in the field of marketing (e.g. consumer's attribution of pricing, advertising etc.) (Silvera & Laufer, 2005). Attribution research has examined consumer's causal inferences for a variety of outcomes (Folkes, 1988):

- Inferences about the consumer's own behavior or the behavior of other persons.
- Inferences about a product's or service's success or failure.
- Inferences about a communicator's endorsement of a product or service.

Summarizing, attribution is of specific interest in the context of CPFs because, 1) attribution is not only linked to purchase outcomes but also to service and product failures (i.e. *failure attribution*) (Folkes, 1984), and 2) research shows that attribution significantly influences various post-purchase behaviors such as complaining, redress seeking, word of mouth, expectancy change, satisfaction and future intentions (Oliver, 1996). As such, failure attribution is an important process through which interaction problems can become CPFs. In the following section, research focused specifically on failure attribution is discussed.

3.3.2 Overview of failure attribution research

Following the definition of the attribution process by Oliver (1996), the failure attribution process (including the special case of a lack of attributional processing) is a mediating phenomenon between observations of product functioning and a number of post purchase behaviors. Failure attributions occur whenever the consumer is primed to reflect on the origins of an outcome. This implies that failures may not be processed if the failure was not significant to the consumer (e.g. a scratch on the backside of an LCD TV) or if the failure was expected (e.g. mobile phone signal loss in a tunnel) (Oliver, 1996). In general, anything unusual which stimulates a person's attention to the failure will bring on causal search. According to Oliver (1996), the most prominent causal agent for attribution is disconfirmation of expectations as shown in Figure 3.6. In the following sections, the antecedents, consequences and measurements of failure attribution are discussed.

Antecedents of failure attribution

According to Kelley and Michela (1980), three types of antecedents for attributions can be defined:

- <u>Motivations</u>: individuals need to be motivated to expend the cognitive effort necessary to determine the cause of an outcome (O'Malley Jr., 1996) and can suffer from motivational biases that can lead to self-serving and false consensus attributions (Folkes, 1988).
- <u>Information</u>: consensus, consistency over time and modality, and distinctiveness influence whether people attribute an outcome to the person, the stimulus, or the situation (Folkes, 1988). In other words, consumers need information or knowledge to determine the cause of the outcome (O'Malley Jr., 1996).

- Prior beliefs: consumers' pre-existing hypotheses, suppositions and expectations can influence the type of attributions made by consumers (Folkes, 1988).

The most important reason why the investigation of consumer knowledge as an antecedent of failure attribution in the context of complex technical products is interesting can be shown by discussing the theoretical distinction in attribution theory between causes and reasons (Oliver, 1996). Causes are agents that are capable of bringing out an event or outcome. Their impact can be direct or indirect and even imperceptible to the receiver. Because of the potential for many causes to be indirect and imperceptible, consumers are unlikely to be aware of such influences and hence may attribute effects to other explanations that are consistent with consumers' existing knowledge. Alternatively, reasons may have little basis of scientific fact but will make perfect sense to the individual and this "wrong" attribution of the perceived failure can still result in a complaint when the failure is perceived to be caused by the product. Reasons are explanatory accounts by the consumer at the level of a consumer's understanding. They may correspond correctly to causes if the consumer is sophisticated in the knowledge category or has otherwise gained knowledge of what caused a particular event (e.g. the notification of the cause of delays by the Dutch railway company).

As discussed in Chapter 1, in the case of complex technical products consumers are less and less aware of and trained in the technical functioning of a product. On the other hand, the opportunity for and the potential span of CPFs becomes increasingly larger, either caused by faults, by the environment or the by consumer him/herself. Consequently, it is very interesting to investigate how the consumer's knowledge affects attribution of different failures in this context.

Failure attribution measurements

Weiner's attribution framework (Weiner, 1985; Weiner, 1986) is the most widely used model to explain and measure differences in failure attribution. This framework is used because it is widely applied in various research fields (Silvera & Laufer, 2005) and because the framework allows the classification of larger sets of attributions within a smaller number of meaningful and actionable categories (Oliver, 1996). The framework consists of the following attribution dimensions (Weiner, 1985; Weiner, 1986):

- Locus: failures can be attributed internally to something within the person or externally to something outside the person.
- Controllability: the cause of the failure can be perceived to be modified by the actor (controllable) or modified by an external agent (uncontrollable). This dimension often interacts with locus (i.e. controllability is undefined unless locus is assigned while uncontrollable failures do not require locus to be defined)
- Stability: the cause of the failure can be perceived as permanent (stable) or temporary (unstable).

In literature not many standardized measurements for failure attribution can be found. Qualitatively, failure attribution can be measured by using open response questions and subsequently coding of the responses into different categories (Oliver, 1996). The most widely used and validated standardized scale is Russell's causal dimension scale (Russell, 1982; Russell, 1987). This scale uses nine items to measure the three attribution dimensions of Weiner's framework. Oliver (1996) adjusted this scale to counterbalance the abstract nature of the statements formulated by Russell (1987).

Consequences of failure attribution

Research on the consequences of attribution (Folkes, 1984; Folkes, 1988, Oliver, 1996) shows that the consumer's perceptions of the causal dimensions of Weiner's framework affect the consumer's expectations of redress for the product failure, anger at the firm and intention to repurchase from the firm:

- Locus influences beliefs about who should solve a failure; failures attributed internally should be resolved by consumers while externally attributed failures should be solved by the firm involved.
- Controllability influences the consumer's desire to hurt the firm's business after perceived product failure such that when firms are perceived to have control over the cause of the failure consumers express more anger than when the firm is perceived to have lack of control over the perceived failure.
- Stability influences expectancies such that stable causes for a failure lead to more confidence that the same failure will recur than do unstable causes (such as the weather).

Although from the above can be concluded that locus is of most importance with respect to predicting CPFs, it is important to note that it is important to measure and analyze all of the above dimensions in order to avoid misleading conclusions when linking failure attributions to complaint behavior (Folkes, 1988).

3.3.3 Conclusion

This section investigated the underlying constructs through which interaction problems result in CPFs. Based on the general consumer processing model of Oliver (1996) both expectation-disconfirmation and failure attribution were identified as relevant mediating variables. It was argued that failure attribution is of specific interest for this research project. Although attribution theory is well-founded in consumer behavior research, few papers can be found that specifically address the antecedents and consequences of failure attribution (Silvera & Laufer, 2005; Weiner, 2000).

Furthermore, although many studies of attribution of service failures and subsequent failure recovery can be found (Smith, Bolton & Wagner, (1999); Harris, Mohr & Bernhardt, 2006; Ma, 2007) few studies have investigated how consumers arrive at attributions of product failures (Folkes, 1988; Silvera & Laufer, 2005; Weiner, 2000). Consequently, it is interesting to further investigate how failure attribution measurement can help to better understand consumer perception of technological product failures in the context of this research project.

3.4 Conceptual research framework and research questions

This section presents the conceptual research framework that is used in the remainder of this dissertation. This framework is based upon the initial research model presented in Figure 2.6 in Chapter 2, adjusted for the insights gained from the additional literature review discussed in this chapter. Cognitive processing of interaction problems in terms of expectation disconfirmation and failure attribution are included as important mediating variables that affect the propagation of interaction problems to CPFs. Furthermore, in the conceptual framework the propagation of product development faults to interaction problems is mediated through *usage behavior* which reflects different aspects of consumer behavior in terms of effectiveness, efficiency and usage patterns (see Hornbæk (2006) for a complete overview of measurements of usage behavior). In other words, differences in consumer knowledge are hypothesized to affect different stages of the fault-complaint propagation: 1) through differences in usage behavior when encountering product development faults, which is expected to result in different interaction problems, and 2) through differences in cognitive processing of encountered interaction problems, which is expected to result in different CPFs. Both relations will be investigated for objective as well as subjective measurements of consumer knowledge which is new compared to previous research on related topics where most often only the effect of one or two consumer knowledge constructs is taken into account and where only few studies have investigated how consumers arrive at attributions of failures in complex CE.

The conceptual research framework presented in Figure 3.7 consists of these hypothesized relations between the consumer knowledge constructs and the propagation of product development faults to CPFs through differences in usage behavior and differences in cognitive processing of encountered interaction problems. Furthermore, the framework contains possible moderating effects of both other consumer characteristics (e.g. demographics) and the usage environment in which the consumer-product interaction takes place. As discussed in section 2.2, in terms of cognitive processing of encountered interaction problems, this dissertation only further investigates the effect of consumer knowledge on failure attribution. Furthermore, the propagation of CPFs to the consumer's response to these CPFs and the effect of consumer knowledge on this response are not further investigated in this dissertation as they are out of the scope of this dissertation.

Please note that the relations between the different consumer knowledge constructs and the other moderating variables on the one hand, and usage behavior and failure attribution on the other hand, are not represented separately. This is because of clarity reasons and because, due to time constraints, only a subset of these relations can be investigated. In each of the following chapters, the hypothesized relations between the independent and dependent variables are further specified for each empirical study conducted.

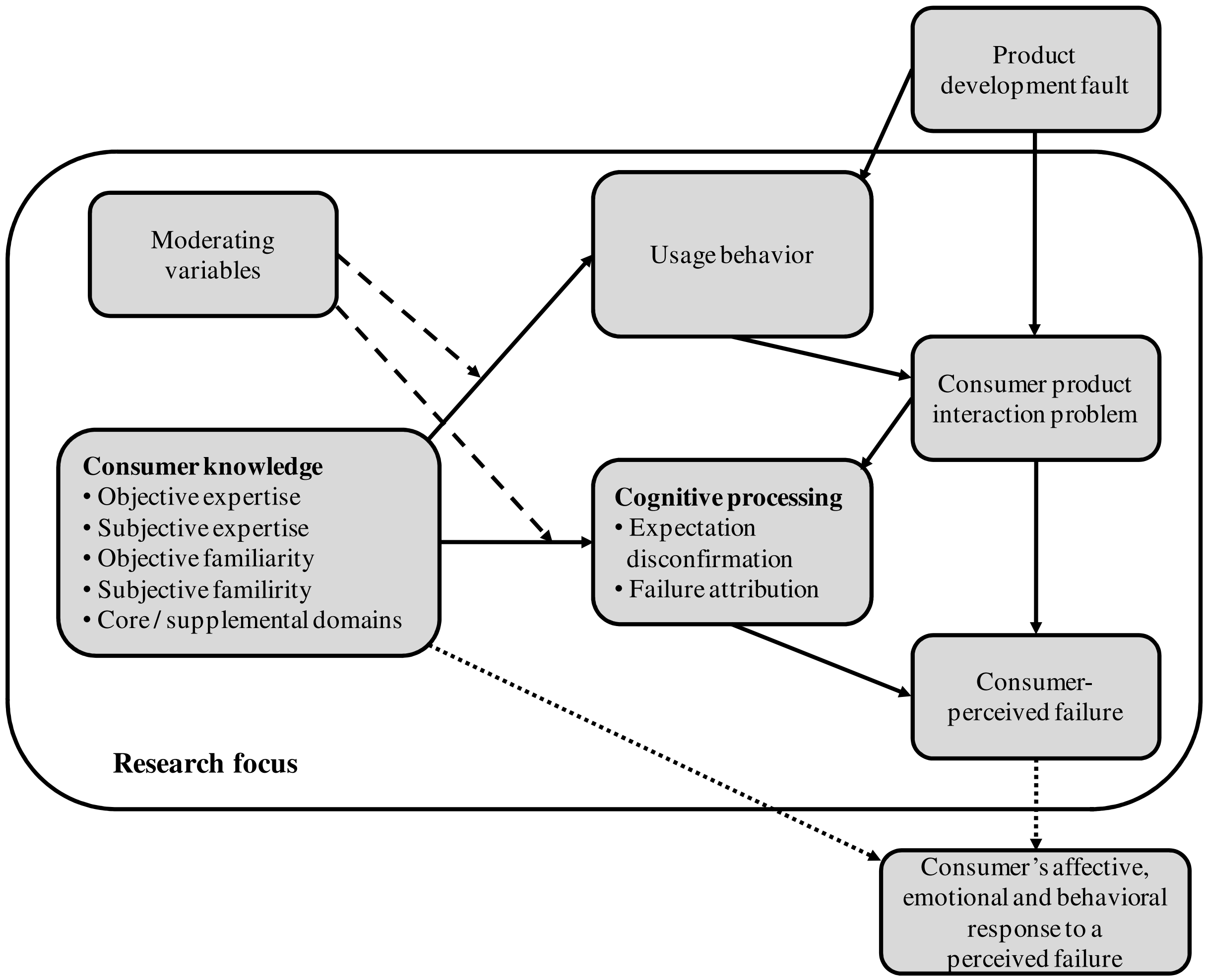


Figure 3.7 Conceptual research framework

Combined with the general aim of this research project defined in section 1.3, the hypothesized relations defined in the conceptual framework result in the following main research question addressed in this dissertation:

How does consumer knowledge affect usage behavior and failure attribution of consumer electronics?

This research question will be answered by dividing it into several sub questions. First of all, the results of the literature review on consumer knowledge constructs show that consumer knowledge measurements are context dependent and should therefore be specifically tailored to the goal of each empirical study. Since no up-to-date and ready-to-use consumer knowledge constructs are available for CE and DTV systems in particular, the first sub research question which needs to be answered is:

1. *How can consumers be differentiated on knowledge of consumer electronics?*

The empirical studies used to answer this sub research question are discussed in Chapter 4 and part of Chapter 7. To fully understand the effect of consumer knowledge on both usage behavior and failure attribution, the second part of the main research question is answered in two parts:

2. *How does consumer knowledge affect usage behavior of consumer electronics?*

3. *How does consumer knowledge affect attribution of product failures in consumer electronics?*

The empirical study used to answer sub research question two is discussed in Chapter 5 and the empirical studies used to answer sub research question three are discussed in Chapters 6 and 7. Before discussing these empirical studies, the general research approach used for these studies is discussed in the following section.

3.5 Research approach and methodology

In this section, the selection of the research approach used throughout the remainder of this dissertation is discussed. As discussed in Chapters 1 and 2, the research presented in this dissertation uses insights from multiple disciplines. Although the research constructs defined in the conceptual research framework are well-researched within each specific field, the combination of these constructs is hypothesized to give new insights for the Q&R field. This and the context of the TRADER project of which this research project is part of, have implications for the selection of the appropriate research approach. Because of the exploratory nature of this research project, an iterative research approach is used. This implies that the research questions defined in section 3.4 are answered in several iterative steps to be able to further refine the use and measurement of the research variables.

The selection of the research methodology for each step of the iterative process depends on a number of factors (Christiaans, Fraaij, De Graaff & Hendriks, 2004; Robson, 1995; Yin, 1994):

- The type of research question
- The extent of control over behavioral events
- The degree of focus on contemporary versus historical events
- The goal of the research project
- The project's constraints and available resources

First of all, because the research questions stated in section 3.4 cover hypothesized relations between behavioral constructs based on insights from theoretical models, the use of experimental methodology seems most appropriate (Christiaans et al., 2004; Yin, 1994). However, the exploratory goal of this research project and the multidisciplinary approach make a pure experimental approach infeasible and also not desirable. A pure experimental approach requires a high degree of control over behavioral events and requires the experimental manipulation of consumer knowledge to be able to randomly assign participants to experimental groups (Stangor, 1998, p.17). As discussed in section 1.3, the goal of this research is to gain more insight into the relation between the *diversity of consumers using complex CE* and the propagation of product development faults to CPFs and subsequent potential consumer complaints. Consequently, other behavioral research designs than "pure" experiments need to be considered to capture the heterogeneity of consumer groups in practice (Stangor, 1998, p. 17; Goodwin, 2005, p.72).

To answer the first sub research question the use of a descriptive research approach by using surveys seems to be the most appropriate. As discussed by Stangor (1998, p. 12), this approach can be used to answer questions on current states of affairs and can as such give a complete understanding of how a larger population of consumers can be categorized on their knowledge of complex CE. To answer the second and third sub research question the quasi experimental approach is used (Cook & Campbell, 1979). This approach is similar to a normal experiment with independent and dependent variables but participants are not randomly assigned to groups. In a quasi experimental approach one of the possible designs is to assign participants to groups based on demographic variables or on the measurement of an occurring characteristic (e.g. consumer knowledge) (Stangor, 1998, p. 262). In other words, this is a sort of experimental approach with a correlation research design (Stangor, 1998, p. 253). When using this approach it is important to consider that no conclusive causal relations can be drawn such as in the pure experimental approaches (Goodwin, 2005, p. 315). However, the approach can be used to give insight into the strength and direction of the relationship (Stangor, 1998, chapter 9) which suits the multidisciplinary approach and the focus on application of research insights rather than gaining insights into fundamental relationships (such as in consumer behavior or psychology research).

Because of the use of an iterative process, it is important to note that the research questions are not answered sequentially and different methodologies and consumer knowledge constructs are used in the empirical research presented in Chapters 4 through 7. An overview of the empirical research discussed in the following chapters is shown in Table 3.2. In this table it is shown which consumer knowledge constructs and which dependent variables are addressed in each chapter. Furthermore, for each construct it is shown whether its measurements are newly developed or based on previous research. As can be seen in Table 3.2, the following chapter discusses the set-up and results of a survey to differentiate consumers on subjective expertise and familiarity of core and supplemental knowledge domains of a multimedia LCD TV.

Table 3.2 Overview of empirical research

Chapter	4	5	6	7
Research question(s) addressed	Differentiation of consumers on subjective expertise, and objective and subjective familiarity on the core and supplemental knowledge domains of CE	Investigation of the effect of consumer knowledge on product usage behavior	Differentiation of consumer on objective expertise of CE and investigation of the effect of consumer knowledge and failure cause on failure attribution	Investigation of the effect of consumer knowledge and failure impact on failure attribution
Methodology	Paper-based survey	Laboratory quasi-experiment	Web-based quasi- experiments	Laboratory quasi-experiment
Consumer knowledge constructs	• Subjective expertise Based on validated scale Flynn and Goldmith (1999) • Subjective familiarity Newly developed scale based on validated constructs • Objective familiarity Newly developed scale based on validated scale in different contexts • Core and supplemental domains Newly selected domains based on comparable research	See Chapter 4	• Objective expertise Newly developed scale based on constructs from Brucks (1985), Arning and Ziefle (2008) and Cordell (1997) • Subjective expertise Shortened scale based on results of Chapter 4 • Objective familiarity Adjusted scale based on results Chapter 4 • Subjective familiarity Shortened scale based on results Chapter 4	• Objective expertise Adjusted based on results Chapter 6 • Subjective expertise Same as Chapter 6 • Objective familiarity Same as Chapter 6 • Subjective familiarity Same as Chapter 6
Dependent variables	None	• Effectiveness Based on validated measurements Hornbæk (2006) and ISO 9241-11 (1998) • Efficiency Based on validated measurements from Hornbæk (2006) and ISO 9241-11 (1998) • Usage patterns Based on validated process mining measurements (Van der Aalst et al., 2007)	• Attribution dimensions Adjusted scale from Russell (1982; 1987) • Number and type of attributed causes Newly developed scale • Perceived picture quality Newly developed scale • Perceived failure impact Adjusted scale from De Visser (2008)	• Attribution dimensions Adjusted based on results Chapter 6 • Number and type of attributed causes Adjured based on results Chapter 6 • Perceived picture quality Same as Chapter 6 • Perceived failure impact Same as Chapter 6

4 Development and validation of subjective expertise and familiarity measurements of consumer electronics

This chapter describes a paper-based survey to investigate how and to what extent consumers can be differentiated on consumer knowledge of multimedia LCD TVs. More specifically, the focus is on categorizing consumers on subjective expertise and familiarity on both core (TV) and supplemental (computer) knowledge domains, taking several moderating variables (age, gender and intention-to-use) into account. The resulting consumer classification is used to investigate how consumer knowledge affects product usage behavior discussed in the following chapter.

This chapter is organized as follows. Section 4.1 describes the conceptual framework that underlies this chapter and the following one. Subsequently, in section 4.2 the design of the survey to investigate the differentiation of consumers on subjective expertise and familiarity of LCD TVs is discussed. Section 4.3 reports on the results of this survey and discusses the reliability and validity of the consumer knowledge measurements developed for this study. Finally, this chapter concludes with a discussion of the results and limitations of this study in section 4.4.

4.1 Conceptual framework

This section discusses the research variables and the subsequent formulation of the conceptual framework, which serves as a basis for Chapters 4 and 5.

As previously discussed in section 3.2.2, consumer knowledge and the selection of suitable consumer knowledge measurements are dependent on both the product category and the consumer behavior variables of interest. Therefore, section 4.1.1 briefly discusses the product category used throughout the empirical studies discussed in Chapters 4 through 8. Subsequently, section 4.1.2 discusses the constructs underlying the measurement of usage behavior as was previously shown in the overall conceptual research framework in Figure 3.7. Based on both the selection of the product category and dependent (usage behavior) variables, the selection of appropriate consumer knowledge measurements for the survey is discussed in section 4.1.3. Furthermore, since research has shown that the effect of consumer knowledge on consumer behavior can only be correctly interpreted by taking context dependent moderating variables into account (Alba & Hutchinson, 1987), section 4.1.4 discusses which control and moderating variables are taken into account. Finally, section 4.1.5 gives an overview of the conceptual framework for this chapter and the following chapter.

4.1.1 Selection of product category

In section 1.1.2 it was discussed that this research project is part of the TRADER project, which focuses on DTV systems as an application domain. DTV systems are TV systems with a complex software architecture (Fischer, 2004; Stroucken et al., 2005; Tekinerdogan et al., 2008). Examples of these TVs are current (high-end) LCD TVs and plasma TVs.

Although nowadays almost every consumer in The Netherlands is, to some extent, familiar with the basic functionality of TVs, taking DTV systems as a case study is still very interesting for this research project. First, as discussed in Example 1.1 in Chapter 1, TVs have changed dramatically from a technological point of view, both in terms of software content and in terms of connectivity requirements. This shift has also resulted in a larger span of potential product development faults (Tekinerdogan et al., 2008). Secondly, DTV systems offer far more functionality than only watching TV programs. They can be used to access the Internet, watch digital photos stored on a solid state storage device and even connect (wireless) to a PC to watch downloaded movie content nowadays. Consequently, these products are highly complex both from a technological and a consumer point of view. It can therefore be assumed that for this product category there is a significant spread of consumer knowledge across a large consumer population (i.e. almost every consumer owns a TV).

4.1.2 Usage behavior of complex CE

In section 3.4, it was discussed that consumer knowledge is expected to influence responses of consumers to any kind of problem they encounter during product usage and as such influence how subsequent interaction problems and which subsequent interaction problems occur in the context of the fault-complaint propagation model.

How consumers use products can be measured in many different ways. In the context of the conceptual research framework shown in Chapter 3 in Figure 3.7, of main interest are those variables which capture how consumers deal with product development faults and related events during usage of complex CE. As such, the goal is not to capture differences in how often a product is used or how many different functions of a product are used by different consumer groups (for example, see the research by Shih and Venkatesh (2004)). As discussed in section 2.3.1, in usability and related research various measurements and techniques can be found that reflect the consumer’s actions when using a system. However, choosing suitable usability measures is difficult and the conclusions of studies in this context depend on the chosen usability measures (Hornbæk, 2006).

The most commonly applied groups of performance measures of usability are effectiveness and efficiency (ISO 9241-11, 1998). Effectiveness is defined as: “the accuracy and completeness with which users achieve specified goals"; and efficiency is defined as: "resources expended in relation to the accuracy and completeness with which users achieve goals" (ISO 9241-11, 1998). In usability literature many different performance measures can be found on both groups (see for example Hornbæk (2006) for an extensive overview). Since these measures are generally accepted both effectiveness and efficiency measures will be used in the laboratory experiment to reflect product usage behavior. Furthermore, in a similar study on the performance of mobile phones these measurements have proven to reflect differences in consumer knowledge (Ziefle, 2002).

However, research also shows that effectiveness and efficiency measures alone do not fully capture the rich, multidimensional and temporal aspects of product usage in consumer tests (Hilbert & Redmiles, 2000; Hornbæk, 2006; Kim & Han, 2008). In other words, these measures can reflect consumer knowledge differences in product usage but, for example, do not give information on different actions taken by consumers to perform a certain task. Consequently, to gain more insight into how consumer knowledge affects differences in product usage behavior, the product usage patterns need to be taken into account as well. For example, Cuomo (1994) shows that sequential analysis techniques such as lag sequential analysis and pattern analysis can help to discover habitual stereotyped patterns of consumer behavior.

Summarizing, differences in consumer knowledge are hypothesized to affect differences in product usage behavior reflected through differences in both standard usability measurements of effectiveness and efficiency and more advanced measures of usage patterns. Which specific usability and usage pattern measures and techniques are used in the laboratory experiment will be further discussed in Chapter 5.

4.1.3 Selection of consumer knowledge measurements

In section 3.2.2 it was discussed that consumer knowledge consists of both familiarity and expertise constructs, which can be measured both subjectively and objectively. Although there is abundant research evaluating the effect of product usage experience or product ownership (i.e. familiarity) on (perceived) product usability (Lazar & Norcio, 2003; Nielsen, 1993; Ziefle, 2002), expertise dimensions as defined in consumer behavior literature (e.g. Alba and Hutchinson (1987)) are often not taken into account. Furthermore, often only objective measurements of familiarity (e.g. usage experience in years or ownership of a product) are used. As a first step, the survey discussed in this chapter and the subsequent laboratory experiment discussed in the following chapter investigate the effect of subjective expertise and objective and subjective familiarity on product usage behavior. The choice is made to focus on these measurements initially (and for now exclude objective expertise) because:

- There are standardized scales available for both subjective expertise and objective measurements of familiarity.

- Subjective expertise reflects differences in an individual's confidence in his/her level of product knowledge, which hypothetically can have an effect on how consumers deal with interaction problems during product usage.
- Objective expertise measurements are highly product and context dependent and are very difficult and time consuming to develop. Since this research is still in an exploratory stage, the choice is made to exclude this measurement at this stage.
- Meta-analysis of consumer knowledge studies[9] has shown that subjective expertise measurements are an adequate surrogate for objective expertise measurements for durable and luxury goods (Carlson et al., 2009).

Furthermore, as discussed in section 3.2.1, Alba and Hutchinson (1987) argue that for different products and tasks different types of knowledge are required to positively affect performance. Since high-end LCD TVs nowadays include Internet browsers and other functionalities emerging from the PC domain, the concept of core and supplemental knowledge domains of Moreau et al. (2001) is used to differentiate consumers on knowledge of LCD TVs. Similar to the original use of this concept to help better predict adoption of discontinuous innovations (Moreau et al., 2001; Saaksjarvi, 2003), it is expected that knowledge on one or more supplemental knowledge domains better facilitates the consumer's understanding of the functionalities of the product under study. In other words, in this first study both knowledge measurements on TVs (core knowledge domain) and on computers (supplemental product domain) are chosen to be taken into account. The expected consumer groups resulting from this classification are shown in Table 4.1. Because it is not yet known whether consumers can be segmented into these four groups for complex CE such as LCD TVs, specific hypotheses on the effect of consumer knowledge in both product domains on usage behavior are formulated in Chapter 5.

Table 4.1 Classification of consumers on core and supplemental product domain knowledge of LCD TVs (adopted from Saaksjarvi (2003)).

		Supplemental knowledge domain (computers)	
		high	**low**
Core domain (televisions)	**high**	Technovators	Core experts
	low	Supplemental experts	Novices

4.1.4 Selection of control variables and moderating variables

To be able to fully understand the relationship between consumer knowledge and product usage behavior several moderating and control variables need to be taken into account (Alba and Hutchinson, 1987; Carlson et al., 2009). Following the definition by Baron and Kenny (1986), a moderator variable is "a qualitative (e.g. gender, race, class) or quantitative (e.g. level of reward) variable that affects the direction and/or strength of the relation between an

[9] Please note that this meta-analysis publication was published after completion of this study

independent or predictor variable and a dependent or criterion variable. Previous consumer knowledge research has shown that age (e.g. Arning and Ziefle (2009)) and gender (e.g. Peracchio and Tybout (1996)) potentially affect the use of consumer knowledge in product usage and evaluation situations and are therefore taken into account as possible moderating variables in this study.

Furthermore, as discussed by Saaksjarvi (2003), the consumer's level of compatibility with a product could potentially affect how consumer knowledge affects usage behavior. Consumers who find a product to be completely incompatible with his/her lifestyle, values and needs, will never voluntarily use this product. For example, the situation can arise that, although a consumer is classified as a "technovator" based on his level of TV and computer knowledge, a lack of motivation to use a high-end complex LCD TV negatively affects his/her willingness to perform product-related tasks. Consequently, the consumer's intention to use high-end LCD TVs needs to be controlled for; consumers with extremely low interest to use this type of product need to be excluded from the laboratory experiment because for these consumers the effect of knowledge could be different or less strong.

Although Saaksjarvi (2003) used the construct of compatibility (of the adoption diffusion model developed by Rogers (2003)) to control for this factor, for this study a more general construct is preferred because compatibility only refers to specific aspects of intention-to-use products. A more abstract and encompassing method to measure intention-to-use can be found in the Unified Theory of Acceptance and Use of Technology (UTAUT) (Venkatesh, Morris, Davis & Davis, 2003). Research has shown that user acceptance of information technology has a strong relationship with the intention to subsequently use this technology (Davis, 1989; Taylor & Todd, 1995; Venkatesh et al., 2003). Although the UTAUT model is designed for use in work environments, its questions can be adapted for use in private situations for CE.

4.1.5 Conclusion

The research variables and its hypothesized relations that were discussed in this chapter are shown in Figure 4.1. Based on this conceptual framework and the classification of consumers on both knowledge domains, hypotheses will be formulated in section 5.1.

In this chapter the first step to answer the first research sub question is taken. The goals of the survey discussed in the remainder of this chapter can therefore be described as follows:

- To set-up a measurement for subjective expertise and familiarity in the TV and computer domain related to complex LCD TVs.
- To validate these measurements in a survey.
- To investigate the differentiation of consumers into segments based upon the measurement scales of subjective expertise and familiarity.
- To select participants for the experiment discussed in Chapter 5.

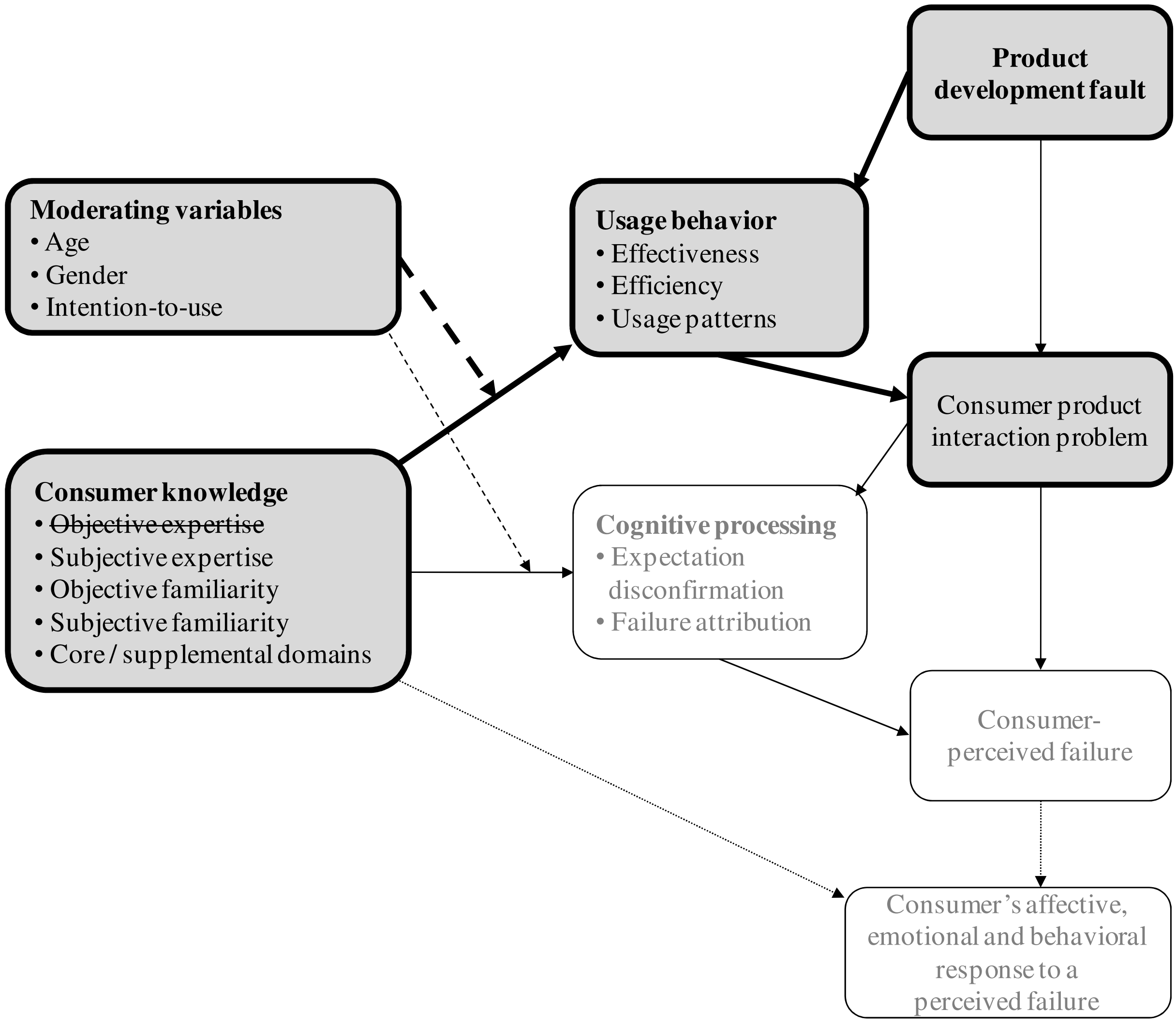


Figure 4.1 Conceptual framework for Chapter 4

4.2 Survey design

In this section the design of the survey, which is used to investigate the differentiation of consumers on subjective expertise and familiarity on both the TV and computer product domain, is discussed. Section 4.2.1 discusses the population and selection of the sample for this study, as well as the data collection method appropriate for this sample. Subsequently, section 4.2.2 discusses the research variables, their relations and the measurement of these variables. Finally, section 4.2.3 discusses the design of the questionnaire by which these variables are measured.

4.2.1 Population, sample and sampling method

Before discussing the set-up of the measurement of the consumer knowledge variables and the design of the questionnaire, the selection of the survey population and sample needs to be discussed. According to Dillman (2000, p. 196) the survey population "consists of all of the units to which one desires to generalize survey results” and “the sample consists of all units of the population that are included in the survey”. The selection of the sample and the appropriate sampling method is of specific importance for the level and type of generalizations that can be drawn for the population (Lohr, 2008, p. 98).

The purpose of this first study is to investigate the categorization of consumers on knowledge on the core and supplemental domains of high-end LCD TVs. Consequently, the population of interest for this study consists of consumers who are willing to use such an LCD TV and meet generally used demographic criteria, such as aged 16 years old or above. Preferably, a heterogeneous population (i.e. in terms of educational level, age, gender etc.) should be used to have a potentially larger range of levels of knowledge on the core and supplemental domain of high-end LCD TVs. As such, the goal of this study is not to generalize the results to the entire population of TV consumers. Moreover, in this context it is very difficult to determine up front which consumers meet the inclusion criteria in the sample such as being interested in using this type of TV. Non-probability sampling is therefore acceptable (Stangor, 1998, p. 103) in which respondents are not selected randomly and the sample does not necessarily statistically represent the population from which the sample is drawn (De Leeuw, Hox & Dillman, 2008, p. 9).

Furthermore, it is important to note that since this research project is conducted in the Netherlands only Dutch respondents are included in this survey and subsequent experiments and surveys discussed in the remaining chapters. One of the consequences of using a heterogeneous population is that the complete questionnaire needs to be written in the Dutch language to prevent nonresponse and measurement error due to not understanding the English language (Lynn, 2008, p. 49). The translation process is further discussed in section 4.3.3.

Based on the discussion above, convenience sampling was used for this survey (Stangor, 1998, p. 104). By using this sampling method it could be ensured that the questionnaire reached different consumer groups and that a sufficient response rate was achieved to be able to recruit a sufficient number of participants for the laboratory experiment discussed in Chapter 5. Although for nonprobability sampling no statistical inferences can be drawn for an entire population and from that perspective no sample size requirements can be formulated, to be able to use the statistical technique factor analysis (Hair, Black, Babin, Anderson & Tatham, 2006) to analyze the validity of the consumer knowledge measurement scales the absolute minimum sample size is 50 respondents or at least five times the number of observations used for factor analysis which is 170 (5 x 34 items, see also section 4.3.3). Taking into account that convenience sampling results in a higher response rate than random sampling, a sample size of 400 was estimated to be sufficient to reach the minimum sample size. A self-administered mail survey was used as the data collection method because of the limited resources required (De Leeuw and Hox, 2008) and because research has shown that reasonable average response rates for this method are reported (De Leeuw, 2008, p. 128).

The distribution of questionnaires for this study started on the 19th of February 2007 and questionnaires returned before the 26th of March 2007 were included for further analysis (i.e. approximately one month response time). Out of the total of 400 questionnaires, 288 were distributed via friends, family, colleagues and students (convenience sample). To reach the required sample size of 400, the remaining 112 questionnaires were sent to postal addresses in the city of Eindhoven in the Netherlands which were randomly selected via a website with a list of all postal codes in the area. Please note that this additional method of respondent recruitment was purely used to exhaust all potential means of recruiting as many respondents as possible and therefore the results cannot be generalized beyond conclusions drawn from a convenience sample.

4.2.2 Research variables

In this section the measurement scales used to measure subjective expertise, objective familiarity, subjective familiarity and intention-to-use are discussed. Unless stated otherwise, the measurements discussed in this section are used for both TVs and computers.

Subjective expertise measurement

In consumer knowledge research, a common measurement to assess subjective expertise is to use a single self-report item (see for example Cordell (1997)) in which consumers are asked to report their perceived level of knowledge on a product or service. However, for single self-report items it is difficult to establish reliability and validity and most of these items were tailored to a specific study. To counter these issues and reduce the possibility of measurement error, a multi-item scale to measure subjective expertise developed by Flynn and Goldsmith (1999) is used. This scale consists of five items reflecting the construct of subjective expertise for which consumers are asked to rate on a seven-point Likert scale whether they agree with the statement or not. This scale has been validated for multiple products, services etc. (Flynn & Goldsmith, 1999). An overview of the items adjusted for the measurement of subjective expertise of TVs is given in Table 4.2.

Table 4.2 Subjective expertise measurement (adjusted from Flynn and Goldsmidt (1999))

Item	Question
1	I know pretty much about televisions.
2	I do not feel very knowledgeable about televisions. (reverse score)
3	Among my circle of friends, I am one of the "experts" on televisions.
4	Compared to most other people, I know less about televisions. (reverse score)
5	When it comes to televisions, I really do not know a lot. (reverse score)

A measure of subjective expertise can be obtained by reversing the scores of the negative items and adding the scores up with those of the positive items. The Dutch subjective expertise measurement scale can be found in Appendix 4.1. For the subjective expertise scale and other scales used in this questionnaire the explicit choice is made to limit the number of response categories to five due to the use of a heterogeneous sample (for example, for older respondents a differentiation of their answer on a seven-point scale could prove to be too difficult to use (see for example Fowler & Cosenza (2008)).

Objective familiarity measurement

In this survey product familiarity is objectively measured by one item measuring the average totality of usage of both televisions and computers (Smith, Caputi et al., 1999; Söderlund, 2002). For both products objective familiarity is measured on a five-point Likert scale. To account for the possibility that a respondent did not use a TV or computer at all, a dichotomous ("yes/no") question was used to filter out respondents with no usage experience of TVs and/or computers (see also section 4.2.3 on questionnaire design issues). These responses were subsequently coded at "0" for objective familiarity. The categories of the TVs objective familiarity measurement are based on the results of a national survey by the Social and Cultural Planning Office of the Netherlands on usage of media by the Dutch population in 2006 (Breedveld et al., 2006). Average use of computers is also reported in this survey but the results only reflect usage during leisure time. Therefore, the categories were estimated to account for usage of computers at both work and at home. The Dutch objective familiarity measurement can be found in Appendix 4.1.

Subjective familiarity measurement

In literature on consumer knowledge, subjective familiarity measurements include scales to measure levels of the consumer's perceived information search on, and usage and ownership of a product (Alba & Hutchinson, 1987; Dodd et al., 2005; Park et al., 1994). However, details on which items are used to measure these variables are not discussed in detail and therefore need to be specifically developed for this study. First of all, ownership and usage are measured by using dichotomous items, which are used to exclude or include subsequent questions on information search and product usage (i.e. if a respondent answered "no" to the television usage question, the respondents is asked to skip subsequent items dealing with television usage). Secondly, subjective measurement of product usage is measured by three items reflecting product usage in daily life and product usage compared to friends and family. Finally, information search is divided into three sub scales reflecting information search during product purchase (e.g. "If I consider to buy a television, I consult multiple information sources"[10]), information search during product use (e.g. "I often seek for information on the use of my television on the Internet") and information search in general (e.g. "I regularly talk to my friends and colleagues about new developments in television products"). All items were measured on five-point Likert scales similar to those used for subjective expertise

[10] Please note that the original subjective familiarity measurement items were developed in the Dutch language and that the items formulated in English in this section are translated by the researcher for explanatory purpose only. Use of these items in research in the English language requires a more thorough translation process (e.g. see Behling and Law (2000)).

measurement. An overview of the subjective familiarity questions (in Dutch) is given in Appendix 4.1.

An overview of the main constructs (besides objective familiarity which has to be analyzed separately due to the use of a different type of measurement than for the other constructs) and underlying constructs discussed above is shown in Figure 4.2. The different levels of underlying constructs will serve as a basis for the statistical analysis of the results of the survey discussed in section 4.3.

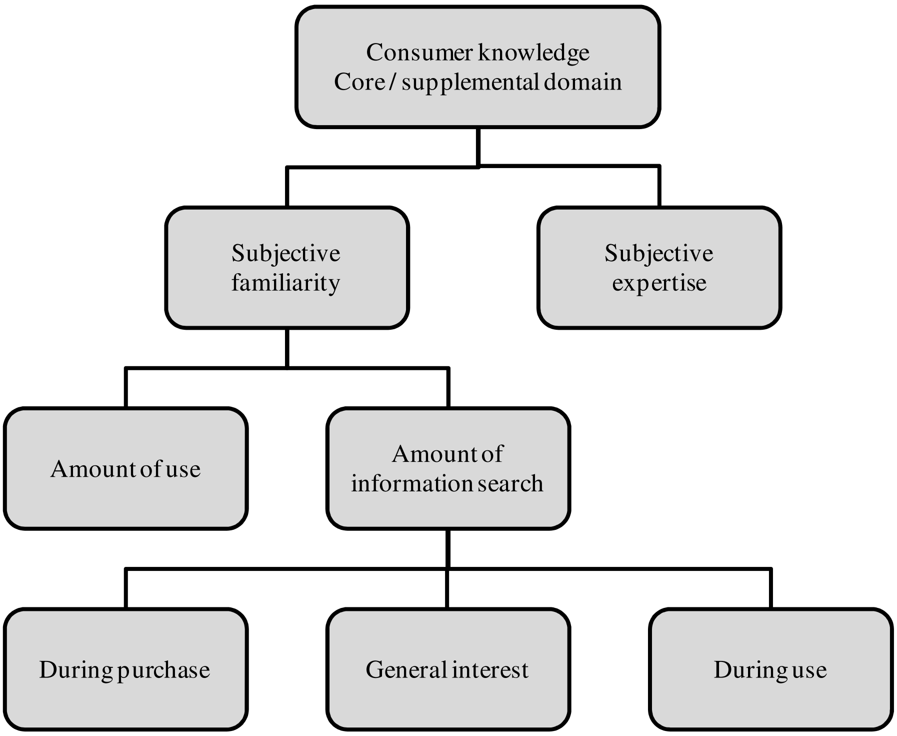


Figure 4.2 Underlying constructs of the questionnaire

Intention-to-use

As discussed in section 4.1.4, the UTAUT model developed by Venkastesh et al. (2003) is used to control for respondents who are not motivated to use high-end LCD TVs and therefore should be excluded from selection for the laboratory experiment discussed in Chapter 5. The original UTAUT model consists of the following four constructs (Venkatesh et al., 2003):

- Performance expectancy: "the degree to which an individual believes that using the system will help him or her to attain gains in job performance".
- Effort expectancy: "the degree of ease associated with the use of the system".
- Social influence: "the degree to which an individual perceives that important others believe he or she should use the new system".
- Facilitating conditions: "the degree to which an individual believes that an organizational and technical infrastructure exists to support use of the system".

From the definition of the constructs can be seen that these constructs are designed to measure intention-to-use information technology in work environments. To apply these concepts to the context of the use of CE in home environments several adjustments need to be made. First, the construct "social influence" is removed from the scale because it does not reflect the context of usage of CE in home environments. Secondly, one item of the performance expectancy scale is removed because it is only valid for work environments ("If I use the system, I will increase my chances of getting a raise"). Finally, several words are changed to match the context to usage of multimedia LCD TVs in the home environment. The adjusted constructs and its items (in Dutch) are shown in Appendix 4.1.

4.2.3 Questionnaire design

To conclude the section on the design of the survey, this section discusses the most important aspects of the design of the questionnaire and cover letter. An overview of the items to measure the research variables is shown in Appendix 4.1

Question and response categories wording and order
Based upon the research variables discussed in section 4.2.2, the final questionnaire consisted of nine different parts in the following fixed order:

1. Introduction to questionnaire
2. Measurement of subjective expertise of TVs
3. Measurement of familiarity with TVs
4. Measurement of subjective expertise of computers
5. Measurement of familiarity with computers
6. Introduction to functionality of multimedia LCD TVs
7. Measurement of Intention-to-use multimedia LCD TVs
8. Personal information (age, gender and educational level and contact information for the experiment and price draw[11])
9. Willingness to participate in the laboratory experiment discussed in Chapter 5[12]

Besides following the general guidelines on questionnaire, question and item construction (Dillman, 2000; Fowler & Cosenza, 2008; Schwarz, Knäuper, Oyserman & Stich, 2008) which will not be discussed in further detail, special attention was paid to two aspects regarding the design of the questions. Several potential order effects were taken into account when ordering these subjects in the final questionnaire (Dillman, 2000, p. 89). First, the measurement of intention-to-use LCD TVs was put *after* the measurement of subjective expertise of LCD TVs because there is a potential of carryover effects of the intention-to-use

[11] Respondents were informed that the contact information was not required when a person wanted to remain anonymous but did want to participate in the survey.
[12] Respondents were given the explicit choice to receive more information about the experiment or not. It was made clear that not willing to participate in the experiment did not affect the participation in the prize draw among the survey respondents.

measurement of *complex* multimedia TVs to the level of expertise a respondent perceives to have about TVs in general. Secondly, personal information and willingness to participate in the experiment were put at the end of the questionnaire because these more personal questions could reduce the response rate when stated at the beginning of the questionnaire. Finally, it is important to consider that it is common to self-administered questionnaires that respondents are the locus of control and they complete the questions without involvement of the researcher in the question–answer process as during personal interviews (De Leeuw & Hox, 2008, p. 261). Consequently, on the cover page of the questionnaire a random person of the family was requested to fill in the questionnaire provided his/her age is at least 16 years old. Furthermore, the questionnaire was designed as such that participants were guided through each section and control questions (i.e. dichotomous items measuring TV and computer usage) were included to allow respondents to skip parts of the questionnaire when s/he, for example, did not own a computer (but still could use a computer in an office environment).

Translation of questionnaire items

All the questionnaire items that were originally formulated in English (i.e. subjective expertise and intention-to-use items of the UTAUT model) were translated into Dutch by using the parallel blind technique (Behling & Law, 2000, p. 23). This method has as an advantage that it can be done faster and allows for more control by the researcher than more traditional translation processes (Behling & Law, 2000, p. 23). Furthermore, since the researchers were fluent in both Dutch and English, the major drawback of this method (lack of source language transparency) is countered. The translation process from English to Dutch was conducted by two persons (an official translator and a person who has been living alternatively in both the Netherlands and the United Kingdom for several years). Both translations were compared by the researcher and a number of discrepancies were resolved by choosing the (perceived) most optimal solution.

Methods to improve response rate

To improve the response rate (besides the advantage of the use of a convenience sample) within the limited resources of the research project, several methods were used of which the most important are (Dillman, 2000; Lynn, 2008):

- Ensuring confidentiality of information provided.
- Use of the university and research project name and logo to emphasize the importance and professionalism of the research project.
- Price draw of €150,- in gift vouchers among the respondents who returned a completely filled in questionnaire as an incentive to return the questionnaire.
- Free response envelope to lower the effort of returning the questionnaire.
- Email address on the cover letter and on the questionnaire to contact the researchers.

Pilot survey

Before the questionnaires were distributed among the target sample, a small pilot survey with five participants was conducted to test the understanding of the cover letter, formulation of items and response categories, question order and complete survey procedure. Based on the comments of the participants, small changes were made in the wording used in the cover letter and in the description of the multimedia LCD TVs described before participants are asked to respond to the statements on intention-to-use.

4.3 Survey results

In this section, the results of the survey are discussed. First, section 4.3.1 discusses the response rate and gives an overview of the characteristics of the respondents included for further analysis. Subsequently, section 4.3.2 discusses the set-up of the statistical analysis of the survey data. In section 4.3.3 the validity and reliability of the consumer knowledge measurements and the intention-to-use measurement are discussed. Finally, this section concludes with a discussion of the classification of the survey respondents on their knowledge on the core and supplemental domain of high-end LCD TVs in section 4.3.4.

4.3.1 Survey response rate and respondent characteristics

In total 240 questionnaires were returned within the defined data collection period. This resulted in a response rate of 60.0%, which is consistent with earlier research findings of mail survey response rated as reported in De Leeuw (2008, p. 128). Out of the 240 returned questionnaires 16 were excluded due to missing answers, which left 224 questionnaires remaining for further analysis. An overview of the characteristics of the respondents is shown in Table 4.3[13].

Table 4.3 Overview of respondent characteristics in terms of age, educational level and gender

Age	n	%	Educational level	n	%	Gender	n	%
< 21 years	18	8.0	Low	31	13.8	Male	138	61.6
21 – 30 years	54	24.1	Medium	95	42.4	Female	86	38.4
31 – 40 years	33	14.7	High	98	43.8			
41 – 50 years	59	26.3						
51 – 60 years	37	16.7						
61 – 70 years	16	7.1						
71 – 80 years	6	2.7						
80 > years	1	0.4						
Total	**224**	**100.0**		**224**	**100.0**		**224**	**100.0**

The results show a reasonable distribution among the age categories; the age of the respondents varied from 16 to 81 years old. The majority of the respondents were male

[13] Please note that no significant differences were found in the characteristics of the respondents recruited via the convenience sample and the respondents recruited via random mail.

(61.6%) and medium to highly educated (86.2%). Possible causes for these effects are the use of a convenience sample (i.e. use of colleagues and students on a technical university) and the possibility that certain consumer groups are more interested in the topic of the questionnaire than other groups and therefore more easily fill in and return the questionnaire (Dillman, 2000, chapter 5). From the 224 questionnaires remaining for further analysis, 33 were filled in anonymous and 60 respondents indicated that they wanted to receive more information on participation in the laboratory experiment discussed in Chapter 5.

4.3.2 Set-up of the data analysis

The goals of this study, as stated in section 4.1.5, were 1) to set-up and validate the measurements of subjective expertise and familiarity in the TV and computer domain related to multimedia LCD TVs; and 2) to investigate the differentiation of consumers into segments based upon their level of knowledge on these product domains. This section discusses the set-up of the data analysis and tests the assumptions underlying this analysis to assess the (construct) validity and reliability of the constructs used in the questionnaire as shown in Figure 4.2. In this context reliability refers to the extent to which the measurements are free from random error (Stangor, 1998, p. 82) while construct validity refers to the extent to which a measured variable actually measures the construct that it is designed to assess (Stangor, 1998, p. 86).

Assessment of scale validity and reliability

To investigate the validity of the consumer knowledge measurements used in the questionnaire, factor analysis is used. According to Hair et al. (2006, p. 104), factor analysis is an interdependence technique that can be used to define the underlying structure among variables in the analysis. In this context, factor analysis can thus be used to investigate whether the four consumer knowledge constructs (i.e. subjective expertise and subjective familiarity on both TVs and computers) as defined in Figure 4.2 are reflected in the survey data. However, from Figure 4.2 it can be seen that subjective expertise is a single construct without underlying constructs while subjective familiarity consists of two lower level constructs that are amount of use and amount of information search. Furthermore, the amount of information search is hypothesized to consist of three lower level constructs. Consequently, there are several layers of constructs that should be taken into account in the factor analysis. To be able to show whether the data reflects these layers in the subjective familiarity constructs, layered factor analyses need to be performed. Objective familiarity measurements will be treated as separate variables and are not included in this factor analysis because these measurements do not involve multiple item scales and are measured with different response categories (i.e. a sixth category was added to account for respondents with no usage experience of TVs or computers). After factor analysis, overall scale reliability and validity are further assessed through computation of Cronbach's alpha and by assessing convergent and discriminant validity of the scales (Stangor, 1998, p. 86-87).

Check of assumptions underlying factor analysis

The first stage of performing factor analysis is to select a specific factor analysis method to use based upon the objective of the data analysis. Since the objective of this factor analysis is to identify latent dimensions of consumer knowledge underlying the items used in the questionnaire, an R factor analysis can be used (Hair et al., 2006, p. 107). Furthermore, before performing factor analysis, several conceptual and statistical assumptions need to be verified (Hair et al., 2006, chapter 3).

With respect to conceptual issues, no independent and dependent variables are mixed in a single factor analysis, the sample is homogeneous with respect to the underlying factor structure and the sample size of 224 exceeds the minimum sample size requirement of 170 respondents (minimum sample size 5 x 34 items = 170 respondents). Concerning the statistical issues, the correlation matrix shows a substantial number of correlations bigger than the cut-off point 0.3 and the anti-image correlation matrix also shows that there are no partial correlations with a value larger than 0.7, which indicates that "true" factors are present in the data. Furthermore, after removal of the variable UseTel_2[14] due to a low Measure of Sampling Adequacy (MSA) value, the MSA for all individual variables, the overall MSA shown in Table 4.4 and the results of a Bartlett's test of sphericity shown in Table 4.4, indicate that all the other variables are suitable for further analysis (Hair et al., 2006, p. 114).

Table 4.4 MSA and Bartlett's test for the initial factor solution

Kaiser-Meyer-Olkin Measure of Sampling Adequacy		**0.873**
Bartlett's Test of Sphericity	Approx. Chi-Square	6276.167
	df	561
	Sig.	.000

Summarizing, based on the assumptions check, the data of this study is suitable for factor analysis. Before proceeding with the discussion of the factor analysis results in the following section, first the factor extraction and rotation method need to be selected. For factor analysis there are two factor extraction methods: common factor analysis and component factor analysis (Hair et al., 2006, p. 117). Although the results of the two factor extraction methods do not differ much when the number of variables exceeds 30, for this study component analysis is used because the primary goal of this analysis is to reduce the data into the smallest number of meaningful factors. Furthermore, factor rotation is required to allow for a more meaningful representation of the questionnaire's constructs. There are two factor rotation methods: orthogonal and oblique rotation methods (Hair et al., 2006, p. 126-127). Since research has shown that consumer knowledge constructs are correlated (Carlson et al., 2009), an oblique rotation method will be used which allows for correlations between constructs (Hair et al., 2006, p. 127).

[14] A description of the abbreviations of questionnaire items can be found in Appendix 4.1

4.3.3 Questionnaire validation

Basic factor analysis of consumer knowledge constructs

Before performing a layered factor analysis in which the number of components extracted from the data is fixed to a certain number based upon the theoretical model defined in Figure 4.2, first a basic factor analysis is performed to investigate whether the data already represent the four consumer knowledge constructs. In this factor analysis there is no fixed number of components to be extracted. According to Hair et al. (2006, p. 128), with a sample size of approximately 200, variables with factor loadings below 0.40 are not statistically significant and are therefore a candidate for exclusion. After three subsequent iterations the following variables were excluded due to significant cross loadings on multiple components and/or due to insignificant factor loadings:

- Subjective expertise of computers: SubExCom_2 and SubExCom_5
- All SearchUTel and SearchUCom items.

Especially the removal of all items reflecting information search during product use is surprising.

The final factor solution for this factor analysis is shown in Appendix 4.2. From this factor solution can be seen that the variables are not grouped together according to the predefined constructs. The general information search items for both product domains and the subjective expertise and amount of use items for computers are grouped together. Consequently, it can be concluded that the collected data does not directly represent the four consumer knowledge constructs and that a layered factor analysis is needed to further interpret these results.

First layer factor analysis of consumer knowledge constructs

The first level analysis was performed with the number of factors to be extracted equal to the number of all individual constructs for both knowledge domains. In total there are 10 individual constructs (5 for each knowledge domain) as shown by the dashed boxes in Figure 4.3.

Similar to the basic factor analysis presented above, the initial solution showed significant cross loadings for the information search during use constructs for both domains Furthermore, a non-significant factor loading for UseTel_3 was observed. After removing these variables there are eight constructs remaining for subsequent analysis. In the second iteration the remaining items are grouped together according to the predefined constructs as shown in the pattern matrix in Table 4.5.

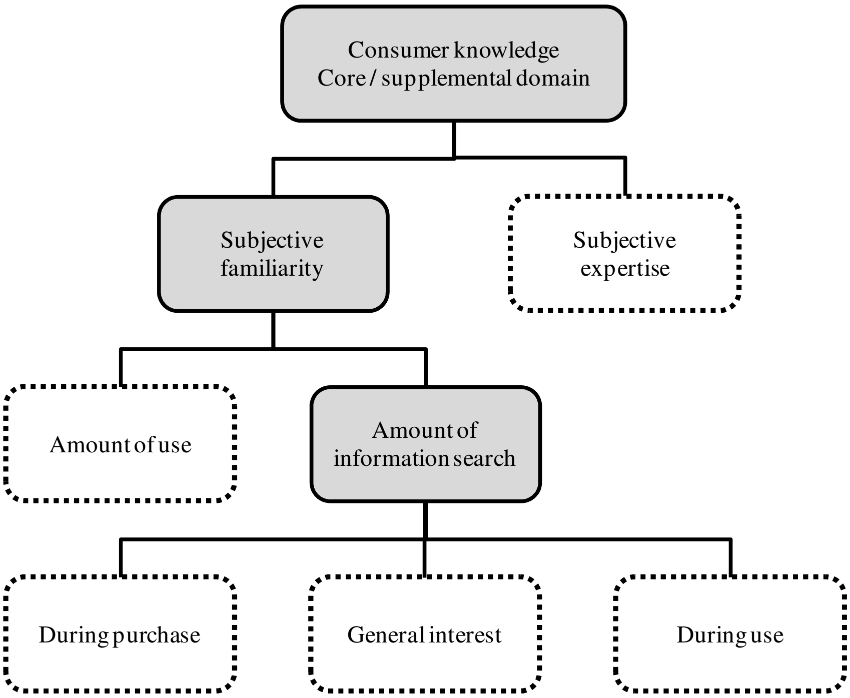


Figure 4.3 First layer of constructs to be tested in factor analysis

Second layer factor analysis of consumer knowledge constructs

The second layer factor analysis was performed by using the results of the first factor analysis as input. In other words, the factor scores resulting from the first level analysis was treated as data in this analysis (for example, general information search on computers was summated to one score based on the individual factor scores for component eight shown in Table 4.5 above). It is now hypothesized that all factors which represent the amount of information search will be grouped together for each product domain. Consequently, the number of factors to be extracted is now six (i.e. three constructs for both the core and supplemental knowledge domain), as shown by the dashed boxes in Figure 4.4.

After excluding several insignificant variables the factor solution in the pattern matrix did not show the expected pattern as shown in this figure because the amount of information search variables cannot be grouped together in one component. Consequently, this factor solution cannot be used for further analysis.

Table 4.5 Pattern matrix of the first layer factor analysis

	Components							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
SubExTel_1				0.708				
SubExTel_2				0.695				
SubExTel_3				0.633				
SubExTel_4				0.867				
SubExTel_5				0.815				
SearchPTel_1		0.931						
SearchPTel_2		0.881						
SearchPTel_3		0.883						
UseTel_2					0.918			
UseTel_4					0.906			
SearchInfoTel_1							-0.807	
SearchInfoTel_2							-0.790	
SearchInfoTel_3							-0.784	
SubExCom_1	0.749							
SubExCom_2	0.775							
SubExCom_3	0.743							
SubExCom_4	0.684							
SubExCom_5	0.766							
SearchPCom_1						-0.874		
SearchPCom_2						-0.835		
SearchPCom_3						-0.908		
UseCom_2			-0.860					
UseCom_3			-0.943					
UseCom_4			-0.872					
SearchInfoCom_1								-0.831
SearchInfoCom_2								-0.828
SearchInfoCom_3								-0.871

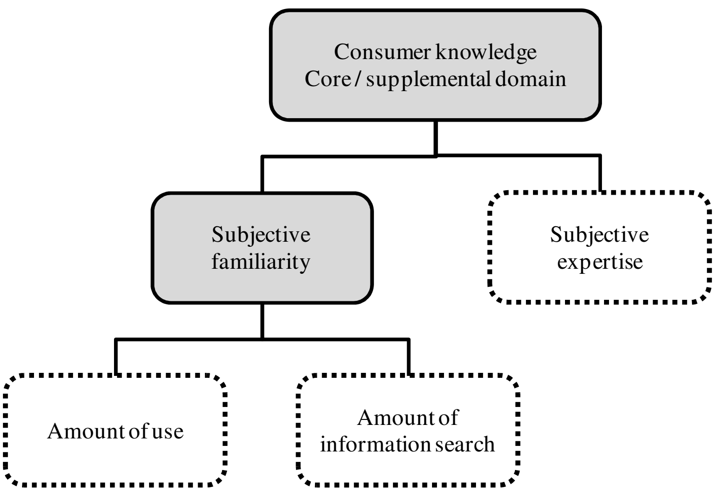


Figure 4.4 Second layer of constructs to be tested in factor analysis

Third layer factor analysis of consumer knowledge constructs

Finally, a third layer factor analysis was performed by grouping the results of the first layer factor analysis into four main constructs as shown by the dashed boxes in Figure 4.5.

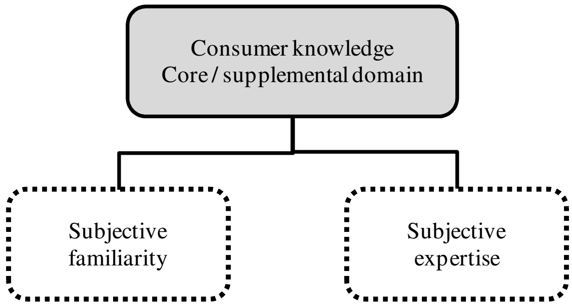


Figure 4.5 Third layer of constructs to be tested in factor analysis

In this instance the number of factors to be extracted is equal to four (two constructs for both the core and supplemental knowledge domain). However, similar to the second layer factor analysis it was not possible to group the constructs underlying subjective familiarity into one factor. It can therefore be concluded that the first layer factor analysis results presented in Table 4.5 represent the final constructs and that there are no deeper layers of constructs behind the developed subjective familiarity construct. In other words, the subjective familiarity constructs as defined in Figure 4.2 are not unidimensional.

Factor analysis of intention-to-use construct

The final factor analysis performed is a factor analysis of the intention-to-use construct. For this construct the same type of factor analysis is used as discussed for the consumer knowledge constructs. The MSA values and the results of the Bartlett's test of sphericity are shown in Appendix 4.3. Due to insignificant factor loadings and significant cross loadings the items measuring facilitating conditions were removed from further analysis. It can be concluded that although these items were adjusted to measure facilitating conditions for use of CE in the home environment, this adjusted construct is not valid in this context. For the remaining two constructs the pattern matrix shown in Appendix 4.3 groups the items together similar to the original UTAUT model by Venkastesh et al. (2003). Consequently, these two constructs are used to represent intention-to-use in further analyses.

Reliability and validity of subjective expertise and familiarity scales

After having analyzed the composition of the constructs underlying the questionnaire, the overall reliability and validity of the remaining constructs is discussed. For the analysis of scale reliability Cronbach's alpha, item-to-total correlations and inter-item correlations are computed. In order to conclude that a scale is reliable, all item-to-total correlations must exceed 0.50, all inter-item correlations must exceed 0.30 and Cronbach's alpha should preferably exceed 0.70 (Hair et al., 2006, p. 137). From the results shown in Table 4.6 can be concluded that all remaining constructs and its items meet these requirements and the constructs are therefore reliable. This table also lists all the valid items that were used for the subsequent scale measurements of the consumer knowledge constructs.

Table 4.6 Overview of reliability and validity of final constructs.

Construct	Valid items	Cronbach's Alpha	Inter-item correlations	Item-total correlations
Subjective expertise televisions	SubExTel_1, SubExTel_2, SubExTel_3, SubExTel_4, SubExTel_5	0.909	0.5 – 0.7	0.68 – 0.85
Amount of information search for televisions during purchase	SearchPTel_1, SearchPTel_2, SearchPTel_3	0.924	0.7 – 0.85	0.80 – 0.90
Amount of use of televisions (subjective)	UseTel_2, UseTel_4	0.809	0.65	0.68
Amount of information search for televisions	SearchInfoTel_1, SearchInfoTel_2, SearchInfoTel_3	0.908	0.7 – 0.8	0.79 – 0.85
Subjective expertise computers	SubExCom_1, SubExCom_2, SubExCom_3, SubExCom_4, SubExCom_5	0.934	0.6 – 0.85	0.74 – 0.89
Amount of information search for computers during purchase	SearchPCom_1, SearchPCom_2, SearchPCom_3	0.947	0.8 – 0.9	0.88 – 0.91
Amount of use of computers (subjective)	Usecom_2, UseCom_3, UseCom_4	0.920	0.75 – 0.9	0.79 – 0.87
Amount of information search for computers	SearchInfoCom_1, SearchInfoCom_2, SearchInfoCom_3	0.928	0.75 – 0.9	0.80 – 0.90
Intention-to-use: performance expectancy	PerfExpMMTV_1, PerfExpMMTV_2, PerfExpMMTV_3	0.872	0.64 – 0.78	0.69 – 0.80
Intention-to-use: effort expectancy	EffExpMMTV_1, EffExpMMTV_2, EffExpMMTV_3, EffExpMMTV_4	0.879	0.55 – 0.82	0.61 – 0.81

Based on the respondents' score on the individual items for each consumer knowledge construct for each product, a mean score for each construct is computed. An overview of the descriptive statistics for each construct is shown in Table 4.7 on the next page. Although the factor analysis did not group the subjective familiarity constructs together, for simplicity reasons and because of the theoretical basis (see section 4.2.2), the scores on these constructs (i.e. amount of information search during purchase, general information search and amount of use) are averaged to form a single mean subjective familiarity score for each knowledge domain.

By using the mean score of the subjective expertise and familiarity constructs, both convergent and discriminant validity can be discussed by investigating the correlations between these constructs. According to Hair et al. (2006, p. 137), convergent validity assesses the degree to which two measures of the same concept are correlated. Consequently, in this study convergent validity can be assessed by evaluating the correlation between the subjective expertise and familiarity measurements within the same product domain. Although these constructs do not fully measure the same concept, research has shown that overall these knowledge constructs overlap and are therefore significantly correlated (e.g. Carlson et al., 2009; Cordell, 1997). Furthermore, discriminant validity refers to the degree to which two conceptually similar concepts are distinct (Hair et al., 2006). In this context discriminant validity can be assessed by investigating the correlations between similar consumer knowledge constructs of the two different product domains.

Table 4.7 Descriptive statistics for the main questionnaire constructs

Construct	Mean	S.D.	Scale range	Number of items
Subjective expertise televisions	2.58	1.22	1 – 5	5
Subjective familiarity televisions	2.65	0.95	1 – 5	8
Objective familiarity televisions	3.46	1.21	0 – 5	1
Subjective expertise computers	2.67	1.28	1 – 5	5
Subjective familiarity computers	2.89	1.25	1 – 5	9
Objective familiarity computers	3.49	1.67	0 – 5	1
Intention-to-use	3.27	1.03	1 – 5	7

Since normality tests indicate that none of these scores fit a normal distribution, Spearman's rho is used to measure construct correlations (Mendenhall & Sincich, 1994, p. 957; Siegel, 1957). An overview of the correlations of the consumer knowledge constructs for both knowledge domains and the intention-to-use construct is shown in Table 4.8. From the results shown in Table 4.8, several conclusions can be made with respect to the validity of the consumer knowledge constructs. For the computer domain the subjective expertise and both familiarity constructs correlate significantly which is an indication for the presence of convergent validity. For example, similar research by Cordell (1997) on consumer knowledge of photo cameras reported a correlation of r=0.50 between subjective expertise and familiarity. However, for the TV domain can be seen that the results for convergent validity are mixed because subjective expertise correlates significantly with subjective familiarity but not at all with objective familiarity. In other words, the results of this study show that using a general measurement of TV does not seem to relate to higher levels of perceived expertise on TVs or higher levels of perceived familiarity with TVs.

Table 4.8 Correlations (Spearman's rho) of questionnaire constructs, N = 224

	Subjective expertise computers	Subjective familiarity televisions	Subjective familiarity computers	Objective familiarity televisions	Objective familiarity computers	Intention-to-use
Subjective expertise televisions	0.619**	0.443**	0.471**	-0.004	0.238**	0.456**
Subjective expertise computers		0.241**	0.665**	-0.177**	0.496**	0.492**
Subjective familiarity televisions			0.524**	0.243**	0.149*	0.319**
Subjective familiarity computers				-0.096	0.580**	0.445**
Objective familiarity televisions					-0.064	-0.009
Objective familiarity computers						0.274**

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).

Next, the discriminant validity of both the subjective expertise and subjective familiarity constructs is low based on the significant and often reasonably high correlation of these measurements between the different product domains. It seems that, although the items referred to distinctly different product domains, from a consumer perspective the *perceived* level of expertise and *perceived* level of familiarity refer to a more abstract level of confidence in using CE and a more abstract level of interest in CE respectively. The very low and not significant correlation between the objective familiarity constructs of both product domains demonstrates that these constructs have significant discriminant validity.

Besides the correlations between the consumer knowledge constructs, Table 4.8 also shows the correlations of these constructs with the measurement of the intention to use multimedia LCD TVs. From the final column of this table can be seen that, apart from objective familiarity on TVs, higher levels of subjective expertise and familiarity on both product domains relate to higher levels of the intention to use multimedia LCD TVs. In other words, excluding survey respondent with low levels of subjective expertise and/or familiarity from participation in the laboratory experiment could also result in bias towards higher knowledge respondents. Implications for the selection of participants for the experiment are further discussed in Chapter 5.

Finally, the correlation between the consumer knowledge constructs and the moderating variables is assessed. From the results shown in Table 4.9 can be seen that age negatively correlates with all consumer knowledge constructs on the computer domain and subjective expertise on TVs, but positively correlates with TV familiarity. Furthermore, there is a positive correlation between male respondents and higher levels of subjective expertise and familiarity. This confirms that both variables need to be taken into account when investigating the effect of consumer knowledge on product usage behavior in the following chapter.

Table 4.9 Correlations of age and gender with questionnaire constructs, N = 224 (calculated with Spearman's Rho for age and Pearson's correlation for gender)

	Subjective expertise televisions	**Subjective expertise computers**	**Subjective familiarity televisions**	**Subjective familiarity computers**	**Objective familiarity televisions**	**Objective familiarity computers**	**Intention-to-use**
Age	-0.289**	-0.344**	0.171*	-0.108	0.225**	-0.220**	-0.215**
Gender	0.428**	0.377**	0.331**	0.346**	0.024	0.117	0.348**

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).

4.3.4 Categorization of respondents

Based on the participants' score on the consumer knowledge constructs, this section investigates how and to what extent consumers can be segmented according to the following four hypothesized segments defined by Saaksjarvi (2003) (as previously shown in Table 4.1):

- <u>Technovators</u>: high TV knowledge and high computer knowledge
- <u>Supplemental experts</u>: low TV knowledge and high computer knowledge
- <u>Novices</u>: low TV knowledge and low computer knowledge
- <u>Core experts</u>: high TV knowledge and low computer knowledge

To differ between high and low knowledge on a product domain, classification boundaries need to be defined. Because it is commonly accepted in consumer knowledge research to use a split on the mean score to differentiate between low and high knowledge (Söderlund, 2002; Sujan, 1985), this study also uses the mean score to differentiate between low and high scores on each consumer knowledge construct. Scatter plots of the resulting categorization of the survey's respondents on the mean scores (i.e. mean score of each respondent on the valid items shown in Table 4.6) of subjective expertise and subjective familiarity are shown in Figure 4.6 and Figure 4.7 respectively.

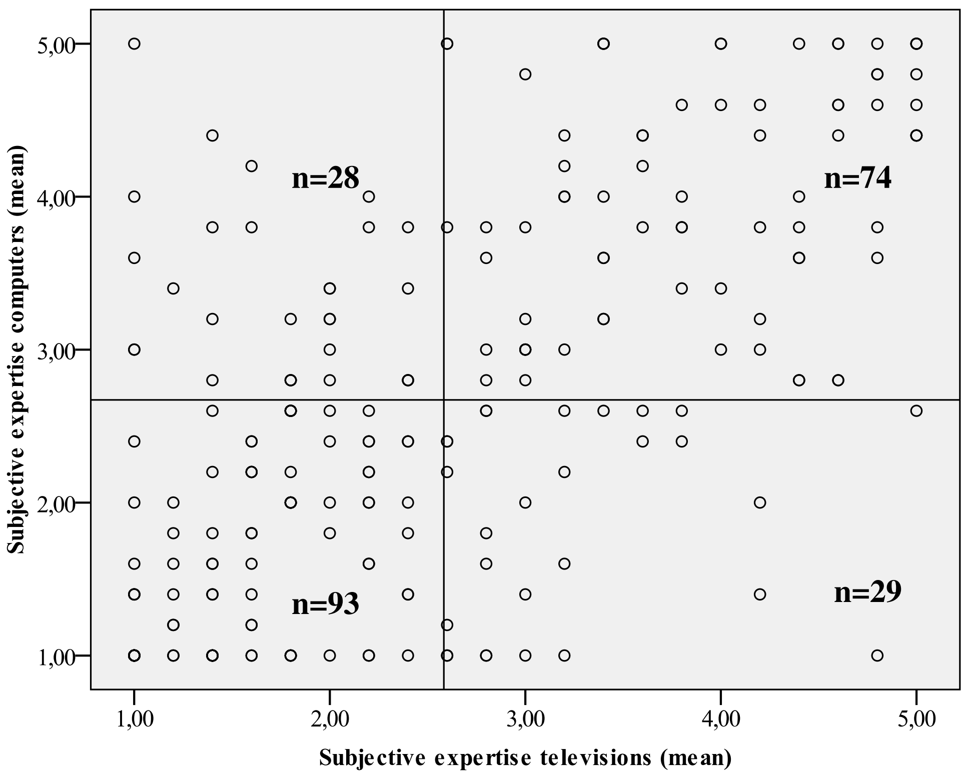


Figure 4.6 Scatter plot of subjective expertise for core (TV) and supplemental (computer) knowledge domains

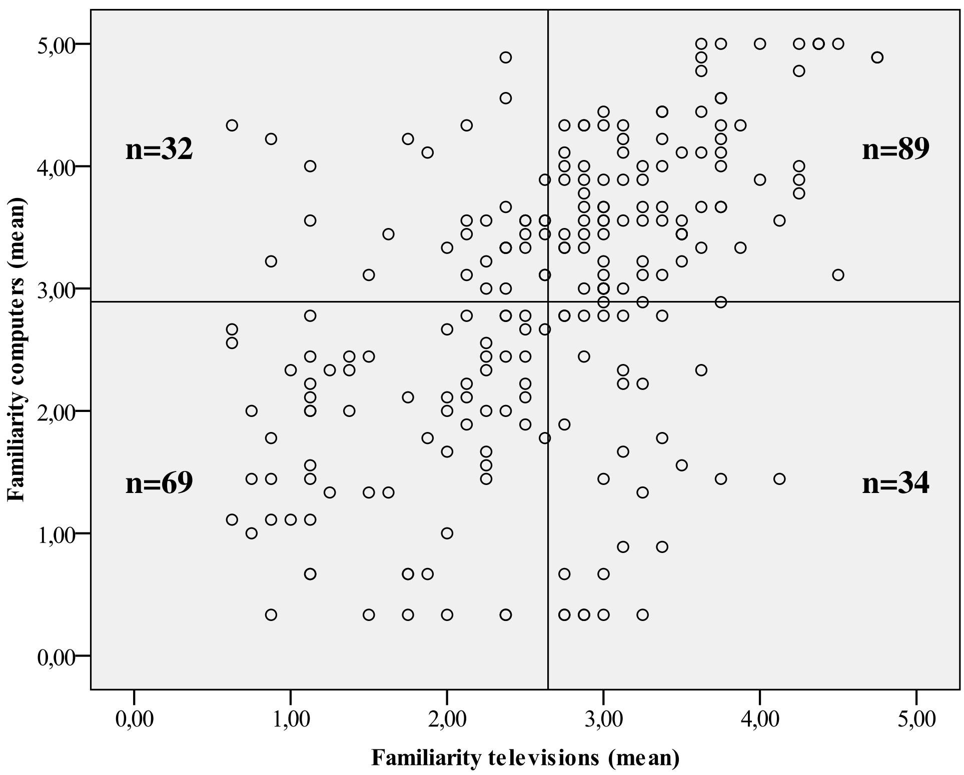


Figure 4.7 Scatter plot of subjective familiarity for core (TV) and supplemental (computer) knowledge domains

For both figures it is important to note that, due to multiple occurrences of the same mean scores, every circle in the graph represents one or more respondents. Since the objective familiarity scale only consisted of one item, no meaningful scatter plot for this construct could be made. An overview of the number of respondents per segment for each of the consumer knowledge constructs is therefore shown in Table 4.10.

Table 4.10 Segmentation of survey respondents (in total number of respondents per segment) on consumer knowledge constructs based on a split of the mean score

			Computers	
		Construct	low	high
Televisions	**Subjective expertise**	high	28	74
		low	93	29
	Subjective familiarity	high	32	89
		low	69	34
	Objective familiarity	high	49	57
		low	49	69

From both scatter plots and the table above can be concluded that, due to earlier discussed reasonably high correlations between subjective expertise and subjective familiarity of the core and supplemental knowledge domains, the resulting distribution of respondents among the four segments of these knowledge constructs is unequal. For the objective familiarity construct an approximately equal distribution can be observed. Although use of other classification boundaries, such as the median score, could result in a slightly more equal distribution of respondents among the segments, the possibilities to segment consumers based on *consumer knowledge domains, as well as subjectively measured consumer knowledge constructs* seem to be limited. Nevertheless, the predictive validity of this segmentation can only be assessed when investigating the effect these different segments have on consumer behavior. This will be further discussed in Chapter 5.

4.4 Conclusion and discussion

The aim of this chapter was threefold. The first aim was to select consumer knowledge constructs that are expected to affect differences in product usage behavior of multimedia LCD TVs. This effect will be investigated in a laboratory experiment discussed in Chapter 5. Based on selected measurements of product usage behavior and on literature on consumer knowledge measurements, subjective expertise and subjective and objective measurements of familiarity were selected. Figure 4.1 shows a conceptual framework of the relation between the selected variables.

The second aim of this chapter was to investigate, develop and validate measurements of subjective expertise and familiarity on the core (TV) and supplemental (computer) knowledge domain of multimedia LCD TVs. Based on literature, subjective expertise (5 items) and objective familiarity (1 item) were measured with validated scales. For subjective familiarity, scales were developed to measure two identified lower level constructs: amount of information search (consisting of the lower level constructs amount of information search during purchase, during use and in general) and (perceived) amount of use. For each (sub) scale three items were developed and pretested (see Appendix 4.1).

The data collected in the survey showed that, besides the subscale measuring amount of information search during use (SearchUse) and one of the items of the amount of use scale for the TV domain (UseTel_2), all the other (sub) scales were valid and reliable. Due to weak validity both the information search during use scales for both product domains and the second item of the amount of use scale for the TV domain (UseTel_2) were removed from the subjective familiarity scale. When reflecting on the information search during use items a possible explanation for the weak validity is that it seems that the items developed to measure this construct do not fully reflect an *amount* in contrast to the other two information search constructs.

The third aim of this chapter was to investigate how and to what extent consumers can be differentiated on the selected consumer knowledge constructs. An interesting and surprising result in this context is that, in contrast to related consumer knowledge research (e.g. Cordell (1997)), objective familiarity of TVs does not correlate with subjective familiarity and subjective expertise of TVs (for the computer domain the constructs do correlate as expected). Although according to Schwarz et al. (2008, p. 27) the use of frequency scales could have resulted in systematic response bias, this is not considered to be of major influence in this study since the constructs for the computer domain do illustrate convergent validity. Nevertheless, for following objective familiarity measurements an open response measurement will be used to prevent this potential bias. A possible explanation for the absence of a significant correlation between subjective familiarity and subjective expertise of televisions could be that the use of TVs is commonly associated with watching TV programs, which refers to a relatively passive interaction (i.e. besides the use of the remote control not much interaction takes place when watching a TV program). Although the item measuring objective familiarity with TVs referred to the use of TVs in general, the item did not specifically refer to or measure different types of more active interaction such as programming TV channels or changing the color settings. In contrast, for computers there is no clear main function and usage can therefore refer to many different types of functions, which could explain the positive correlation with subjective expertise.

The results of the layered factor analyses also showed that, although the items measuring subjective familiarity are valid and reliable, they do not reflect a unidimensional subjective familiarity construct. Although due to simplification reasons the item scores were grouped together to form a single subjective familiarity score, it is not known whether this was the best way to treat the subjective familiarity constructs. Further research on the nature of these familiarity constructs in different product domains is needed to be able to decide how to treat the layered familiarity constructs. This is beyond the scope of this research project. Nevertheless, the results do show that the selected and developed consumer knowledge constructs to a large extent empirically hold and thus can be used to differentiate consumers.

The data collected in the survey showed mixed results for the validity of the use of core and supplemental knowledge domains to differentiate consumers on knowledge of complex CE. The results showed that both the subjective expertise and subjective familiarity scales are significantly correlated which also resulted in an unequal distribution of consumers across the four hypothesized consumer knowledge segments. As discussed in section 4.3.3, it seems that, although the items referred to distinctly different product domains, from a consumer perspective the *perceived* level of expertise and *perceived* level of familiarity refer to a more abstract level of confidence in using CE and a more abstract level of interest in CE respectively. For objective familiarity an approximately equal distribution was observed. Consequently, the results of this study suggest that only objective measurements of consumer knowledge can be used when differentiating consumers on multiple and related consumer knowledge domains. Nevertheless, any segmentation based on a split on a scale measurement is artificial and the predictive validity can only be assessed when investigating whether differences in consumer knowledge on multiple knowledge domains also affect product usage behavior.

Finally, it is important to note that the results of this study depend on the use of a convenience sample and the assumption that the TV and computer domain represent the core and supplemental domain of multimedia LCD TVs. Future research could investigate the effect of using a different sample and different product domains on the classification of consumers in the hypothesized consumer knowledge segments.

5 Evaluating the effect of subjective expertise and objective familiarity on product usage behavior

In Chapter 4 it was investigated how and to what extent consumers can be differentiated on knowledge of the core (TV) and supplemental (computer) domain of multimedia LCD TVs. Familiarity and subjective expertise measurements of both product domains were validated in a survey. Based on the resulting categorization of consumers on these measurements, this chapter investigates how both subjective expertise and objective familiarity differences affect product usage behavior when consumers are asked to perform tasks with varying levels of complexity in an experiment with a multimedia LCD TV.

This chapter is organized as follows. Section 5.1 further elaborates on the conceptual framework of this study based on the results of the survey discussed in Chapter 4. Participant selection criteria are defined and hypotheses for each of the dependent variables are formulated. Subsequently, in section 5.2 the design of the experiment to test these hypotheses is discussed. Section 5.3 reports on the results of this experiment, assessing both the effect of subjective expertise and objective familiarity. Finally, this chapter concludes with a discussion of the results and limitations of this study in section 5.4.

5.1 Conceptual framework and hypotheses

This section further discusses the conceptual framework for the investigation of the effect of consumer knowledge on product usage behavior. Section 5.1.1 further elaborates on the goals and overall design of this study in relation to the overall research framework developed in Chapter 3. Subsequently, in section 5.1.2 an adjusted conceptual research framework and hypotheses are presented.

5.1.1 Goals and overall design of the experiment

In the overall conceptual research framework shown in Chapter 3 in Figure 3.7, differences in consumer knowledge are hypothesized to affect the fault-complaint propagation through differences in product usage behavior and cognitive processing of subsequent interaction problems. In this context, the specific goal of this chapter is to investigate how consumer knowledge differences on the core and supplemental domain affect the propagation of faults to potential interaction problems through differences in usage behavior. As stated in section 3.5, a quasi-experimental research approach is used which allows for non-random assignment of participants to groups. Before formulating hypotheses for this experiment, several specific choices concerning the design of the experiment are discussed.

First of all, in contrast to the teletext experiment discussed in section 3.1, for this experiment no deliberate (software) faults were implemented in the LCD TV because:

- For reasons of internal validity it is important to control for possible confounding variables that would limit the interpretation of the results of a study (Goodwin, 2005, p. 165). Since deliberately introduced software faults (which have to have a certain level of severity to be noticed by all participants) also possibly trigger cognitive consumption processing (i.e. failure attribution and related variables discussed in section 3.3.1), this factor needs to be controlled for when investigating differences in usage behavior.
- Reproducing realistic and controllable software failures in DTV systems for usage in real-life experiments is practically unfeasible because of the complexity of such systems and their dependence upon input from the environment (DTV system experts could only create reproducible and controllable failures in a DTV system when completely simulating a DTV system and not allowing the participants to check or even see cables because of this simulation). In contrast to the exploratory teletext experiment where a CRT TV was used, the remainder of this dissertation focuses on complex CE and DTV systems in particular. Any compromise on product appearance due to technical reasons (e.g. simulating an LCD TV with a monitor) would seriously threat ecological validity when evaluating usage behavior.

Instead of trying to implement specific software faults, task complexity is used as a proxy measure for product development faults. This choice is based on input from experts and the industrial partners in the project. Task complexity impacts both technological complexity of a product and cognitive complexity for consumers.

The use of task complexity as a proxy measure for product development faults allows for a more in-depth investigation of differences in usage behavior without possible confounding factors due to deliberately introduced software faults. Although research on usability has already demonstrated the significant effect of differences in familiarity on product usage behavior for tasks with a varying level of complexity (e.g. see Ziefle (2002)), the goal of this study is to further build upon these findings and contribute to research by: 1) investigating the effect of *different* consumer knowledge constructs for *multiple* knowledge domains; and 2) investigating qualitative and quantitative differences in usage patterns of different consumer knowledge groups.

Secondly, due to the use of a quasi-experimental research design, not all consumer knowledge constructs and control variables defined in section 4.1 can be taken into account in the experiment. Although the selected constructs are correlated, participants with a low subjective expertise on TVs do not necessarily have a low level of subjective or objective familiarity on

TVs. It is therefore not feasible to take into account all constructs *based on a split on the mean value obtained from the survey data* due to the resulting large inequalities in group sample sizes, especially considering the fact that out of the 224 respondents only 60 participants had interest possibly participating in the experiment. Since research has shown that the expertise based component of consumer knowledge is more directly related to behavior than familiarity (Brucks, 1985; Cordell, 1997) and because the teletext experiment already partially investigated the effect of familiarity, differentiation on subjective expertise on both product domains was selected as the main selection criterion. To allow for a comparison with the effect of objective familiarity on usage behavior, allocation to high and low levels of this consumer knowledge construct was recalculated based on a split of the mean score for the participants to the experiment only. However, this naturally also resulted in a slightly different allocation and a small difference compared to the survey results for which specific care must be taken when interpreting the results. Since subjective familiarity measurements were not equally distributed, this construct will not be further taken into account.

5.1.2 Adjusted framework and hypotheses

Based on the discussion on the design of the experiment above, the conceptual research framework for this experiment is further narrowed down as shown in Figure 5.1. As discussed above, subjective familiarity is not taken into account in further analyses and gender cannot be used as a potential moderating variable due to the high correlation with the subjective expertise measurements.

The results of the literature review on consumer knowledge discussed in section 3.2 indicated that, in general, both familiarity and expertise positively affect the ability to perform product-related tasks successfully. Since the subjective expertise measurements for both product domains are strongly correlated and as a results only a very limited group of survey participants could be classified as core or supplemental experts, for subjective expertise no differentiation is made based on product domain. This results in the following hypotheses for the general effect of consumer knowledge on product usage behavior:

Hypothesis H_1: Consumers with higher levels of knowledge, measured as follows:

a) Subjective expertise
b) Objective familiarity of TVs
c) Objective familiarity of computers

perform tasks more effectively than do consumers with lower levels of the same measures of knowledge.

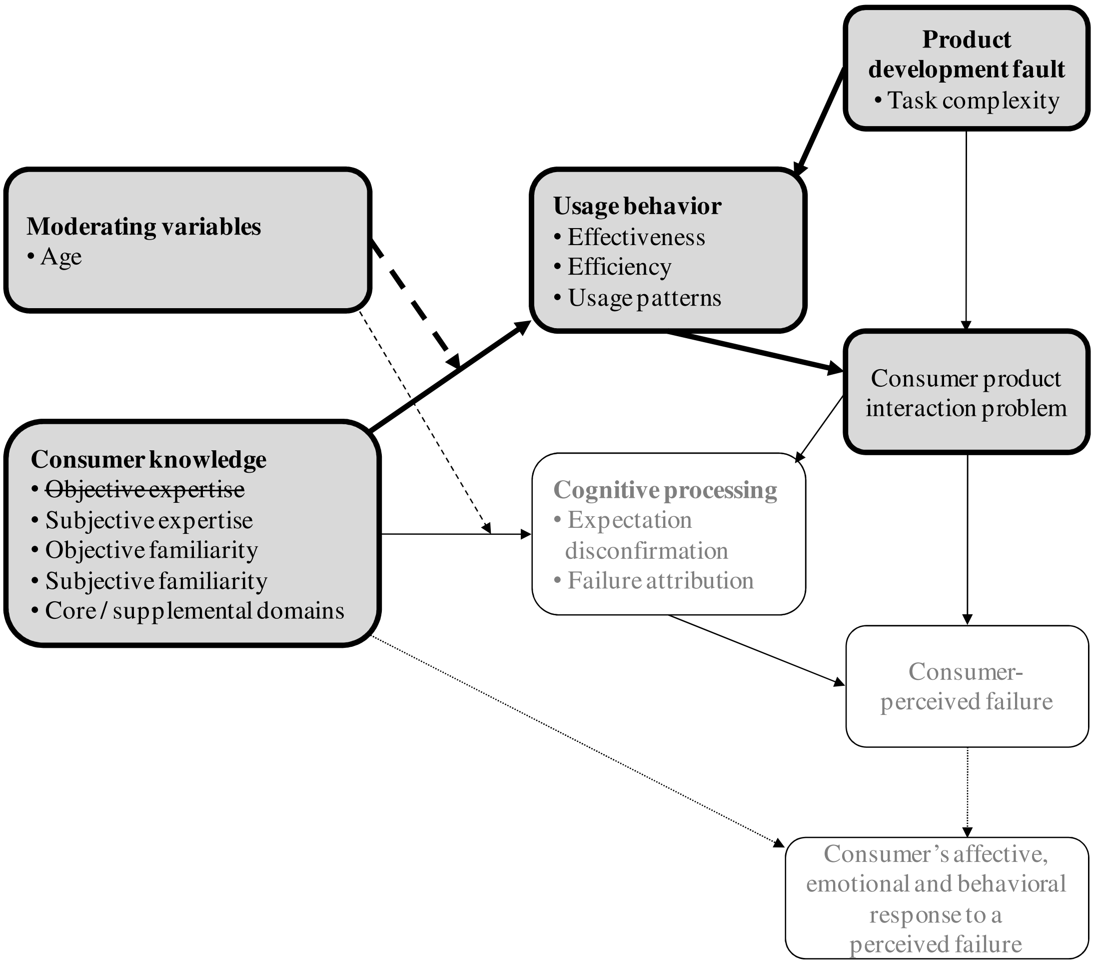


Figure 5.1 Adjusted conceptual research framework to investigate the effect of consumer knowledge on product usage behavior

Hypothesis H_2: Consumers with higher levels of knowledge, measured as follows:

a) Subjective expertise
b) Objective familiarity of TVs
c) Objective familiarity of computers

perform tasks more efficiently than do consumers with lower levels of the same measure of knowledge.

Based on the selection of efficiency measurements, this hypothesis will be subdivided into four separate hypotheses in section 5.2.1.

As discussed above, because research has shown that the expertise based component of consumer knowledge is more directly related to behavior than familiarity (Brucks, 1985; Cordell, 1997), the following hypothesis is stated:

Hypothesis H_3: Differences in subjectively measured expertise stronger relate to differences in product usage behavior than differences in objectively measured familiarity.

Finally, although no differentiation in the hypotheses could be made for the effect of subjective expertise due to the correlation between the measurements for both domains, such a differentiation was made for the effect of objective familiarity. Research has shown that simple repetition of tasks leads to increased performance (Alba & Hutchinson, 1987) and thus can be hypothesized that for tasks related to LCD TVs differences in TV usage experience stronger relate to differences in product usage behavior than differences in computer usage experience. In other words:

Hypothesis H_4: Differences in objective familiarity of TVs have a stronger effect on product usage behavior than differences in objectively familiarity of computers.

5.2 Method

To test the hypotheses formulated in section 5.1.3, a 3 x 2 between-subjects experiment was designed in which selected participants from both the high and low subjective expertise group were asked to perform three different tasks, with a different level of complexity and originating from both the TV and computer domain, on a multimedia LCD TV selected for this experiment. This section describes the set-up of this experiment.

5.2.1 Experimental variables

The independent variables under study are subjective expertise and objective familiarity of the TV and computer domain, and task complexity. Subjective expertise of both TVs and computers was varied on two levels (high and low) based on a split on the mean value of subjective expertise obtained from the results from the survey sample discussed in Chapter 4. Similarly, objective familiarity of both product domains was varied on two levels (high and low) based on a split on the mean value of objective familiarity from the subjects participating in the experiment. The final independent variable was task complexity, which, derived from Ziefle (2002), is defined as the complexity of the menu structure (i.e. number of menu levels and number of distinctly different keys) a participant had to use to complete a task.

As discussed in section 4.1.2, as dependent variables the standard usability measurements of effectiveness and efficiency were used (Hornbæk, 2006; ISO 9241-11, 1998), complemented by qualitative measurements of the usage patterns. Finally, a measurement of satisfaction was included because this variable is part of a standard usability measurement. However, this variable does not relate to the core of the research model tested in this experiment. An overview of all the variables and their parameters and measurements is shown in Table 5.1.

Table 5.1 Overview of dependent variables and accompanying parameters and measurements

Variable	Parameter(s)	Measurement(s)
Effectiveness	Task completion	Dichotomous (completed / not completed)
Efficiency	Time	Total time (in seconds) needed to complete a task
	Steps	Total number of steps in the menus needed to complete a task
	Levelup	Total number of times a participant returns to a higher level in the menu
	Usage patterns	Type and sequence of steps used to complete a task
Satisfaction	ASQ (Lewis, 1991; Lewis, 1995)	Satisfaction on a seven-point Likert scale for: • Ease of completing the task • Amount of time it took to complete the task • The support information (documentation, messages etc.) when completing the task

Effectiveness was measured per task with a dichotomous item referring to the ability to complete the task as specified in the task list. Furthermore, efficiency was quantitatively measured by three parameters: task completion time (in seconds), number of steps in the menu needed to complete the task and number of detour steps (number of returns to a higher level in the menu). Qualitative measurements of efficiency were performed by recording the type and sequence of steps used by participants to complete a certain task. Process mining methods and tools (i.e. ProM) (Van der Aalst et al., 2007) will be used in section 5.3 to identify and analyze the underlying usage patterns extracted from logged actions of the participants. Satisfaction was measured per task by using the After Scenario Questionnaire (ASQ) which addresses three aspects of satisfaction with system usability: ease of task completion, time to complete a task and adequacy of the support information (Lewis, 1991; Lewis, 1995).

Based on the selected measurements for efficiency, hypothesis 2 can be subdivided into the following hypotheses:

Hypothesis H_{2a}: Consumers with higher levels of knowledge, measured as follows:

a) Subjective expertise
b) Objective familiarity of televisions
c) Objective familiarity of computers

need less time to perform tasks than do consumers with lower levels of the same measure of knowledge.

Hypothesis H_{2b}: Consumers with higher levels of knowledge need, measured as follows:

a) Subjective expertise
b) Objective familiarity of televisions
c) Objective familiarity of computers

need less steps to perform tasks than do consumers with lower levels of the same measure of knowledge.

Hypothesis H_{2c}: Consumers with higher levels of knowledge, measured as follows:

a) Subjective expertise
b) Objective familiarity of televisions
c) Objective familiarity of computers

make less detour steps than do consumers with lower levels of the same measure of knowledge.

Hypothesis H_{2d}: Consumers with higher levels of knowledge, measured as follows:

a) Subjective expertise
b) Objective familiarity of televisions
c) Objective familiarity of computers

conform more to the ideal usage pattern than do consumers with lower levels of the same measure of knowledge.

5.2.2 Experimental tasks

All participants were asked to solve three different tasks, each concerning a different functionality of the LCD TV with a different level of complexity. The tasks were selected as such that they covered the more innovative functionalities of a multimedia LCD TV and that they covered both the core and the supplemental knowledge domains for this type of product (i.e. TV and computer domain related tasks). The following tasks were selected for the experiment:

- Dual screen task: simultaneously displaying two specified TV channels on the TV screen.
- Digital picture task: displaying digital pictures, which are stored on a USB stick, on the TV screen.
- Channel switch task: changing the number under which a certain TV channel is stored.

An overview of the task complexity for each of these tasks is shown in Table 5.2.

Table 5.2 Overview of task complexity

	Minimum number of menus	Minimum number of steps
Dual screen	1	4
Digital picture	2	5
Channel switch	3	7

From this table can be seen that the dual screen task is the least complex task and that the switch channel task is the most complex task.

5.2.3 Participants

Out of the 60 survey respondents who indicated that they were willing to take part in the experiment, 29 people (19 male and 10 female, all native Dutch) participated in the experiment[15]. Although this sample size is relatively small, the use of multiple tasks per participant resulted in an acceptable overall sample size that sufficiently meets the requirements for the statistical analyses of the results. All participants received a €20 gift coupon and reimbursement of travel expenditures as compensation for their time and effort.

Based on their mean score on subjective expertise of TVs, the participants were split into a high and low subjective expertise group. Out of the 29 participants, 14 were categorized as high on subjective expertise of TVs (all male) and 15 were categorized as low on subjective expertise of TVs (5 males). The characteristics of both groups based on this differentiation are shown in Table 5.3. An overview of the characteristics based on a similar (due to the correlation of the subjective expertise measurements) although slightly different differentiation on subjective expertise of computers is shown in Appendix 5.4. Please note that due to the quasi-experimental design and the correlation between subjective expertise and gender this resulted in an unequal distribution of males and females among the groups.

Table 5.3 Overview of participant characteristics based on differentiation on subjective expertise of TVs

	High SubExTel (n = 14)			**Low SubExTel (n = 15)**		
	mean	**S.D.**	**range**	**mean**	**S.D.**	**range**
Age	37.86	12.46	22 – 59	48.87	14.90	24 – 67
Intention-to-use	4.48	0.37	3.86 – 5.00	3.22	1.07	1.00 – 4.86
Subjective expertise TVs	4.11	0.90	2.60 – 5.00	1.81	0.44	1.00 – 2.40
Subjective familiarity TVs	3.23	0.93	0.75 – 4.38	2.62	0.94	0.75 – 3.88
Objective familiarity TVs	3.43	1.34	1.00 – 5.00	3.80	1.15	2.00 – 5.00
Subjective expertise computers	4.40	0.73	2.40 – 5.00	1.97	0.56	1.00 – 2.80
Subjective familiarity computers	3.97	0.92	2.00 – 5.00	2.46	1.08	0.33 – 3.56
Objective familiarity computers	4.50	1.29	1.00 – 5.00	3.53	1.60	0.00 – 5.00

[15] Half of the survey respondents who indicated interest in participating in the experiment were eventually not included (1) due to travel distance to the consumer test facility or other practical reasons; and (2) because respondents who scored low on consumer knowledge were also less willing to participate in the experiment while approximately equal sample sizes were required for statistical analyses.

Separate pair wise Mann-Whitney U tests showed significant differences between the two subjective expertise groups for subjective expertise of TVs ($p < 0.001$), subjective expertise of computers ($p < 0.001$) and subjective familiarity with computers ($p < 0.001$). These results confirm that differences on subjective expertise between the two experimental groups are significant and thus can be used for further analysis. Although there is also a significant difference for subjective familiarity with computers, the confounding effect of these differences is limited since the survey results have shown that these constructs are correlated and theoretically they are both part of the same overall consumer knowledge construct.

Because a quasi-experimental research methodology is used, it is important to test for the presence of possibly confounding factors embedded in the two subjective expertise groups. Separate pair wise Mann-Whitney U tests showed significant differences between the two groups for intention-to-use ($p < 0.001$) and age ($p < 0.05$). The potential confounding effect of age is countered for taking this variable into account as a covariate in the statistical analysis (Hair et al., 2006, p. 406). However, because a significant difference for the intention-to-use measurement exists between the two groups and because differences in product acceptance potentially affect usage behavior (as discussed in section 4.1.4), the potential confounding effect needs to be evaluated separately.

5.2.4 Apparatus and materials

The experiment was performed in a consumer test facility on the university campus. This laboratory consists of two rooms, one which resembles a living room in which the test participant performed the test and the other in which the test participant could be observed by the researcher through a one-way mirror. A snapshot of the set-up in the simulated living room is shown in Figure 5.1.

Figure 5.1 Picture of the set-up of the LCD TV in the consumer test facility

For the experiment a 42” HD ready multimedia LCD TV was used with relatively new TV features such as a USB port to display multimedia (e.g. pictures, videos) on the TV screen, support of a wireless connection to a PC and HDMI connectors. At the moment of conducting the experiment (April 2007), this product was only recently introduced on the market. The basic installation settings such as TV channel installation, location and time were already set and these were reset before the start of each experiment. For each experiment the product information sheet, manual, remote control and a USB stick containing several digital pictures were provided.

Each experiment was recorded with two cameras, a non-obtrusive camera mounted on the ceiling to capture the screen of the LCD TV and a camera on a tripod that was used to capture the test participant. Participants were asked to think-aloud during the experiment. For each experiment, two observers were present: one behind the one-way mirror to observe and record the experiment and one in the living room to record the participant’s comments.

5.2.5 Procedure

At the beginning of the experiment the participants were instructed that the goal of the experiment was to evaluate the ease-of-use of the LCD TV. Furthermore, the participants were provided with basic information on the LCD TV (e.g. price, time of market introduction, innovative functionalities) and the participants were instructed that the initial installation was already completed. Subsequently, each participant was asked to read a one page introduction to the experiment (see Appendix 5.1) before starting with the tasks. Each task was explained on a separate page and the participants were asked to complete a task before proceeding with the next one. After each task, the participants were asked to fill in a short questionnaire containing the ASQ items and two control items on previous experience with similar tasks. The task list and task questionnaire (in Dutch) are shown in Appendix 5.2 and 5.3 respectively. The ordering of tasks was randomized to counteract possible learning effects. An overview of the complete experimental procedure is shown in Figure 5.2.

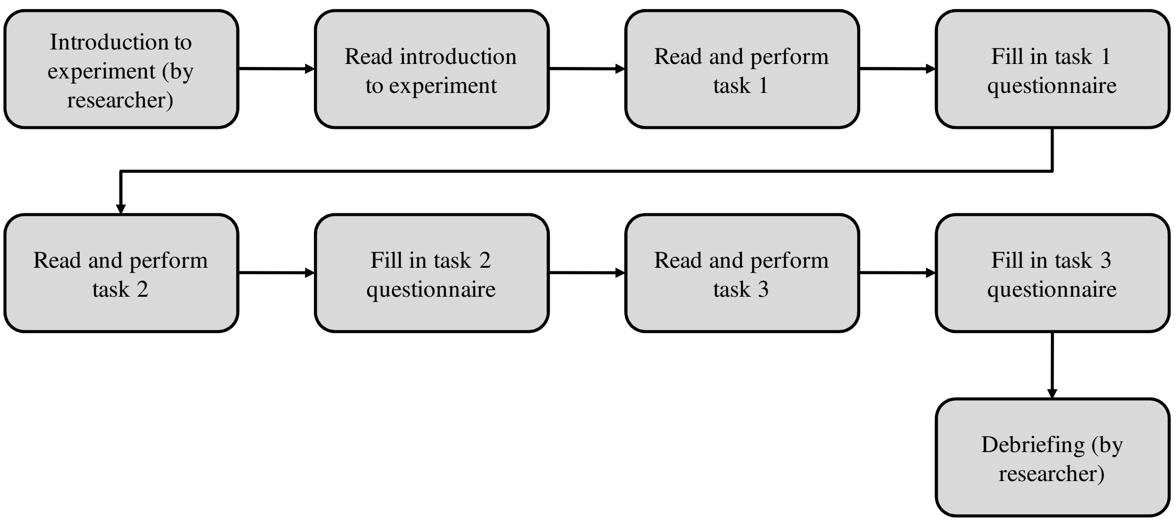


Figure 5.2 Overview of experimental procedure

For those participants who were not able to complete the task because they either decided to quit and/or after a long time could not find the appropriate function, efficiency measurements were excluded from the analysis because these could not be compared with the results of the participants who did complete the task. For the participants who did reach a *perceived* end of the task but did not fulfill the task completion requirements (e.g. programming the channel to a wrong number) efficiency measurements were included (i.e. they did reflect actual usage) but effectiveness was coded as "did not complete the task".

The video data of each experiment was analyzed by logging events and time stamps using Noldus Observer XT software (Noldus). Based on these data, the basic efficiency measurements were calculated. Subsequently, ProM*import* (Günther & Van der Aalst, 2006) was used to convert log files for the process mining analysis in ProM (Van der Aalst et al., 2007). Finally, statistical analyses were performed with SPSS.

The experimental set-up, procedure, measurements and task questionnaire were pre-tested in a pilot experiment with three participants recruited from the group of survey respondents who indicated to be willing to participate in the experiment.

5.3 Results

In this section the results of the experiment are discussed. First, in section 5.3.1 the set-up of the statistical analyses and the analyses of the usage patterns with ProM are discussed. Subsequently, section 5.3.2 discusses the evaluation of the overall effect of the consumer knowledge constructs on the dependent variables. Finally, in sections 5.3.3 – 5.3.5 the results are discussed for each task separately.

5.3.1 Set-up of analyses

Statistical analyses

To investigate the main effect of subjective expertise of TVs and computers on product usage behavior, Multivariate Analysis of Variance (MANOVA) is used. According to Hair et al. (2006, p. 383), MANOVA is a dependence technique that measures the differences for two or more metric dependent variables based on a set of categorical (nonmetric) independent variables (Hair et al., 2006, p. 383). Because the effect of independent variables can be assessed for multiple dependent variables simultaneously, MANOVA allows for control of experiment wide Type I error rate (Hair et al., 2006, p. 400) (as opposed to separate ANOVA analyses).

For this experiment multiple separate MANOVAs (for each consumer knowledge construct separately) with a 2 (high and low subjective expertise) x 3 (task complexity) factorial design are used in which the efficiency measures time, steps and levelup are included as metric dependent variables. As such, subjective expertise is treated as a between-subjects factor and task complexity is treated as a within-subjects factor. To control for the effect of age, this variable is initially included as a covariate. Results of the assumptions check for the MANOVA analysis showed that the data (after variable transformation) sufficiently meet the criteria for performing this analysis. Furthermore, results of a factor and scale reliability analysis show that the ASQ scale to measure satisfaction meets all the criteria for further statistical analysis.

The effect of subjective expertise on the remaining, nonmetric, dependent variables (i.e. satisfaction and task completion) is investigated by using nonparametric separate pair-wise Mann-Whitney U tests. For both the MANOVA and the nonparametric tests the level of significance is set at $p = 0.05$. Results within the less restrictive level of $p = 0.1$ are indicated as marginally significant. Finally, for the MANOVA analyses the significance of the omnibus F-tests were taken from the Pillai values.

ProM analyses

In section 5.2.1 it was discussed that, in addition to standard efficiency measurement, process mining (by using the ProM software) is used to investigate the underlying usage patterns. As discussed by Van der Aalst et al. (2007), the goal of process mining is to discover, monitor and improve real processes (i.e. occurring in reality) by extracting knowledge from event logs. In these event logs occurrence of activities in a process are recorded. While process mining techniques were originally developed for analyses of business processes for the development of information systems (Van der Aalst et al., 2007), they could also be useful for analysis of usage patterns in CE. Process mining can be used to (Van der Aalst et al., 2007):

- Discover new models (e.g. constructing a model that reproduces observed behavior).
- Check the conformance of a model by checking whether the modeled behavior matches the observed behavior in reality.
- Extend an existing model by projecting information from the event logs onto an initial model.

For this experiment both constructing usage pattern models for different levels of subjective expertise and analyzing conformance of usage patterns with the designed ideal usage patterns are of interest. For model discovery several analyses are used:

- <u>Dotted chart analysis</u> (Song & Van der Aalst, 2007): this analysis visualizes the spread of and time between the activities recorded in an event log for each participant.
- <u>Performance sequence diagram analysis</u> (Hornix, 2007): this analysis helps to identify common and rare usage patterns and extreme usage patterns leading to bad task performance.
- <u>Control flow discovery methods</u> (Weijters and Van der Aalst, 2003): by using a so-called heuristic miner differences between high and low levels of subjective expertise can be investigated by creating and subsequently comparing heuristic nets for both groups which visualize common steps taken and for example indicate common "loops" when a participant made a mistake in executing the task.

For conformance analysis log coverage and model fitness measurements are calculated (Rozinat & Van der Aalst, 2008) based upon a comparison of actual usage patterns with "designed" usage patterns deducted from the UI design and the manual (i.e. ideal sequence of steps to perform a task). In this context log coverage refers to a measurement of the match between events in the event log and events specified in the designed usage model (Rozinat & Van der Aalst, 2008). The more often events are logged which do not occur in the designed usage model, the more inefficient a usage pattern probably is. Finally, the measurement of fitness refers to the extent to which the log traces can be associated with valid execution paths specified by the designed usage model (Rozinat & Van der Aalst, 2008). Since both the dotted chart analysis and the conformance analysis refer to a comparison of overall task efficiency, these are discussed in section 5.3.2. The performance sequence diagram analysis and control flow discovery are performed for each task separately and are discussed in sections 5.3.3 – 5.3.5.

5.3.2 Evaluation of overall effect of consumer knowledge

In this section, the overall effect of consumer knowledge on the usage behavior measurements is discussed as well as the results for the evaluation of control variables (effect of task complexity and age) and possible confounding factors (effect of intention-to-use).

Evaluation of confounding effect of intention-to-use

A post-hoc analysis of the potential confounding effect of intention-to-use on usage behavior (using the same factorial design with a split on the mean value of intention-to-use of the survey respondents) did not show a significant effect on the efficiency measurements ($F(3, 66) = 1.581$, $p < 0.3$). Separate pair-wise Mann Whitney U tests confirmed intention-to-use also did not significantly affect both effectiveness and satisfaction. Consequently, it can be concluded that the potential confounding effect of intention-to-use in this experiment was at least limited. Nevertheless, since the survey results discussed in Chapter 4 have shown that this construct is correlated with consumer knowledge, care must be taken when assessing the results of quasi-experimental research.

Effect of subjective expertise and task complexity on efficiency measurements

First of all, the effect of subjective expertise of TVs ($F(3, 66) = 7.773$, $p < 0.001$) and subjective expertise of computers ($F(3, 66) = 9.591$, $p < 0.001$) on the standard efficiency measurements proved to be highly significant. Since adding age as a covariate in the model did not improve both significance and power of the effect of subjective expertise of TVs ($F(3, 65) = 5.310$, $p < 0.01$) and subjective expertise of computers ($F(3, 65) = 5.119$, $p < 0.01$), this variable is removed as a covariate from further analyses (Hair et al., 2006, p. 419). Details of all analyses are shown in separate tables in Appendix 5.5.

Secondly, the highly significant effect of task complexity when using both subjective expertise of TVs ($F(6, 134) = 5.080$, $p < 0.001$) and subjective expertise of computers as an independent variable ($F (6, 134) = 4.678$, $p < 0.001$), confirmed a successful design of the experiment. Further univariate tests for the model using subjective expertise of TVs as independent variables confirmed that task complexity had a significant effect on time ($F (2, 74) = 3.597$, $p < 0.05$), levelup ($F (2, 74) = 9.792$, $p < 0.001$) and number of steps ($F (2, 74) = 16.969$, $p < 0.001$). In other words, overall the least complex task (dual screen) took significantly less time, steps and detour steps to complete than the most complex task (switch channel).

Next, both the subjective expertise of TVs model ($F (6, 134) = 1.380$, $p < 0.3$) and the subjective expertise of computers model ($F (6, 134) = 0.846$, $p < 0.6$) showed that the interaction effect of subjective expertise and task complexity is not significant. Although for the individual measurements the results are not always consistent in accordance with the level of task complexity (as will be discussed next), these results show that for the efficiency measurements the effect of subjective expertise on product usage behavior is not weakened by increasing levels of task complexity.

Since both subjective expertise constructs were highly correlated and the results of the analyses above confirmed that there is an almost equally significant observed effect for both the models tested, in the following only the effect of *subjective expertise of TVs* on the efficiency measurements is discussed. From this discussion analogies can be drawn for the effect of subjective expertise of computers. Results of univariate tests for the effect of subjective expertise of TVs on the efficiency measurements showed a significant effect on time to complete a task ($F (1, 74) = 18.285$, $p < 0.001$) and a marginal significant effect on levelup ($F (1, 74) = 3.304$, $p < 0.1$). However, no significant effect on the number of steps ($F (1, 74) = 1.903$, $p < 0.2$) was observed. Consequently, hypotheses H_{2a} and H_{2c} can be accepted for subjective expertise while H_{2b} needs to be rejected.

Interaction plots of the three (transformed) efficiency measurements are shown in Figure 5.4 – 5.6[16]. From these plots can be seen that there is no interaction for both the dual screen and the switch channel task and the efficiency measurements decrease equally with an increase in both subjective expertise and task complexity.

---

[16] Please note that because the scores of the dependent variables had to be transformed to meet the requirements for the statistical analysis, only these transformed variables can be used in the interaction plots to allow for a proper interpretation of the analyzed interaction effects.

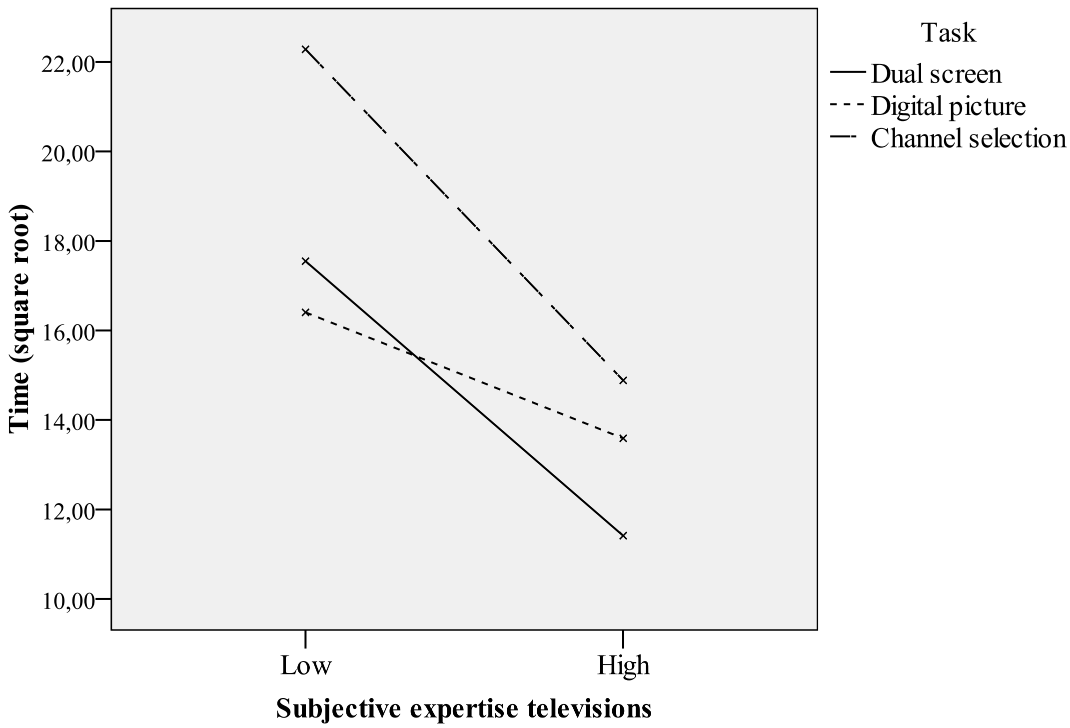


Figure 5.4 Interaction plot of subjective expertise on TVs and task complexity for the time measurement

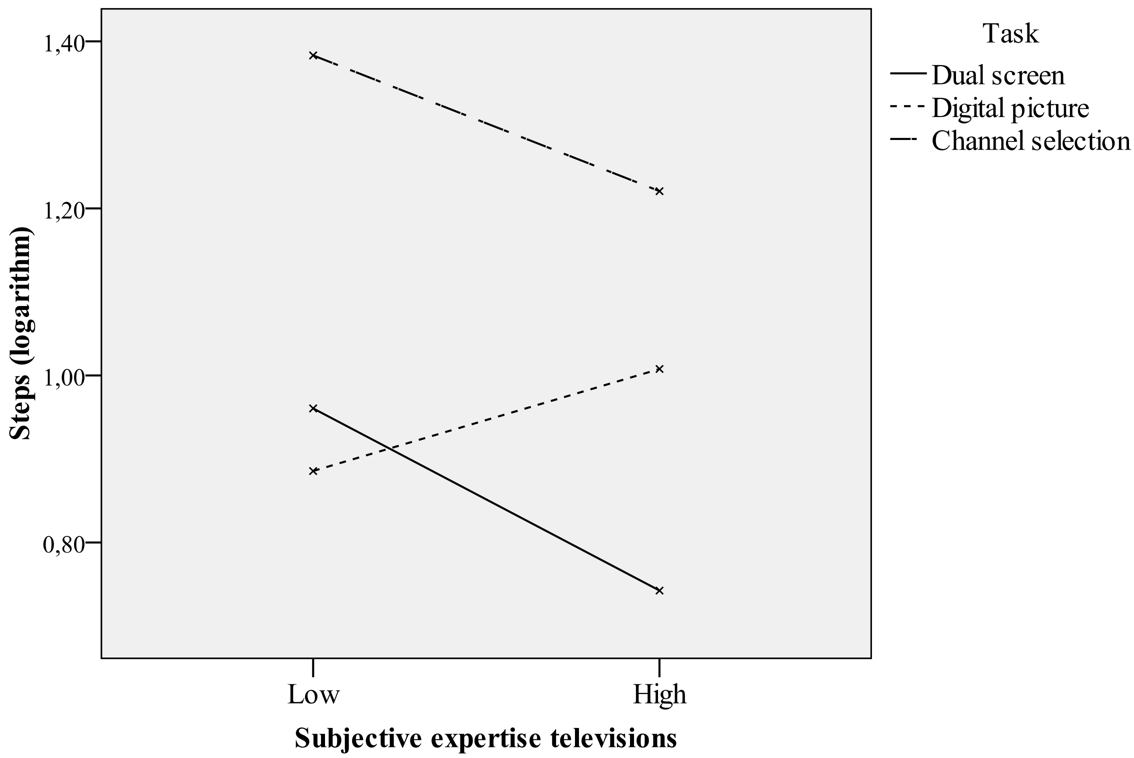


Figure 5.5 Interaction plot of subjective expertise on TVs and task complexity for the number of steps measurement

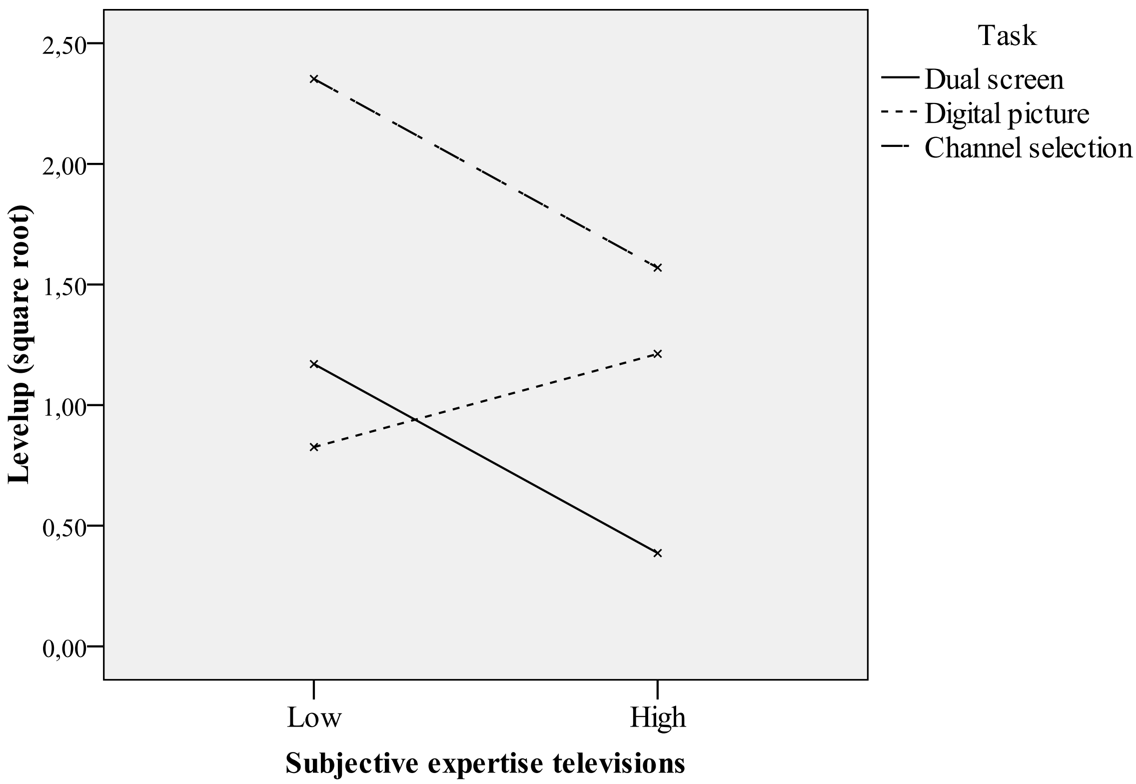


Figure 5.6 Interaction plot of subjective expertise on TVs and task complexity for the levelup measurement

However, the interaction plots clearly show an interaction when looking at the results of the digital picture task. Although this interaction is disordinal, it is not significant and therefore does not negatively influence the interpretation of the main effect of subjective expertise (Hair et al., 2006, p. 420). It does indicate that the results of this task are not consistent with the hypotheses and need to be further investigated in the following sections.

Finally, a dotted chart analysis and conformance analysis were used to further investigate overall differences in usage patterns. The results of the dotted chart analysis are shown in Figures 5.7 and 5.8 on the next page. Please note that the purpose of showing both graphs is only to show the number of and spread of events (each dot in the graph represents one event such as an action performed by a participant) for all the tasks combined and that it is not the purpose to analyze each event separately. The analysis indicates two differences between the subjective expertise groups that support the findings stated for the other efficiency measurements:

- A larger number of logged events for the low subjective expertise group, which indicates that this group used more and different steps to execute a task.
- A larger spread of logged events (i.e. space between the events) across time, which is an indicator for more experienced problems between separate events (e.g. looking for support information in the manual or thinking about which button on the remote control should be used to access a certain function).

Figure 5.7 Dotted chart analysis of low subjective expertise on TVs group (the vertical axis displays different participants while the horizontal axis displays the participants' actions over time)

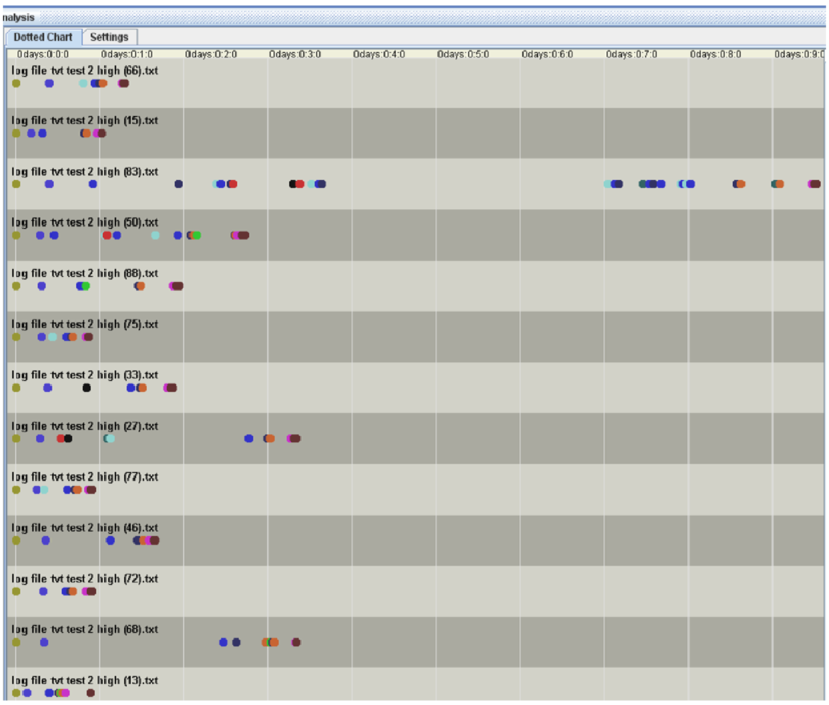


Figure 5.8 Dotted chart analysis of high subjective expertise of TVs group (the vertical axis displays different participants while the horizontal axis displays the participants' actions over time)

For the conformance analysis the designed usage models were created by using the extended product manual and subsequently modeling the process with Petri nets by using Yasper (Yasper, 2009). The designed usage models are shown in Appendix 5.6. When comparing these models with the event logs, the results of the conformance analysis shown in Table 5.4 also support the significant effect of both task complexity and subjective expertise on the standard efficiency measurements. In case of task complexity, the measurements show a decrease in log coverage and level of fitness for increasing levels of task complexity. The results for the log coverage analysis show that the designed usage model covers fewer events than actually occurred when analyzing the usage patterns for the low subjective expertise group for both the dual screen and switch channel task compared the usage patterns for the high subjective expertise group. For example, the designed usage model covers less than 50% of the events logged when analyzing the usage patterns of the low subjective expertise group for the switch channel task. As such this is an indication for inefficient usage patterns or a bad conceptual model of the product's designer on how consumers perform these tasks. Furthermore, the results of the level of fitness measurement show that, for the participants in low subjective expertise group, a slightly lower proportion of log traces fitted into the designed usage model as well (i.e. was a step also specified in the designed usage model) relative to the current position of the process (i.e. the position in the designed usage model is based on past steps). The results shown in Table 5.4 also confirm that for the digital picture task the differences between the subjective expertise groups are in the other direction than hypothesized in section 5.3.1. Consequently, only for the dual screen and switch channel task hypothesis H_{2d} can be accepted and further analysis of the usage patterns is needed for each task separately to draw further conclusions on this hypothesis. These separate analyses are discussed in the following sections.

Table 5.4 Overview of overall log coverage and level of fitness measurements for ` differentiation on subjective expertise

Task	Log coverage		Level of fitness	
	High subjective expertise TVs	**Low subjective expertise TVs**	**High subjective expertise TVs**	**Low subjective expertise TVs**
Dual screen	0.667	0.543	0.972	0.951
Digital picture	0.662	0.813	0.940	0.956
Switch channel	0.529	0.498	0.887	0.787

Effect of subjective expertise and task complexity on effectiveness

An overview of the results of the measurement of task completion is shown in Figure 5.9. From this figure, by comparing the average task completion measurements for each task can be seen that the digital picture task had the highest task completion ratio followed by the dual screen task and the switch channel task. One-way Kruskal-Wallis tests confirmed that there is no significant overall effect of task complexity on effectiveness (χ^2 (2) = 4.663, p < 0.1). In other words, although the digital picture task, by design, required more menus and steps to complete, the task was on average completed more often than the dual screen task. This corresponds with the results of the efficiency measurements discussed above.

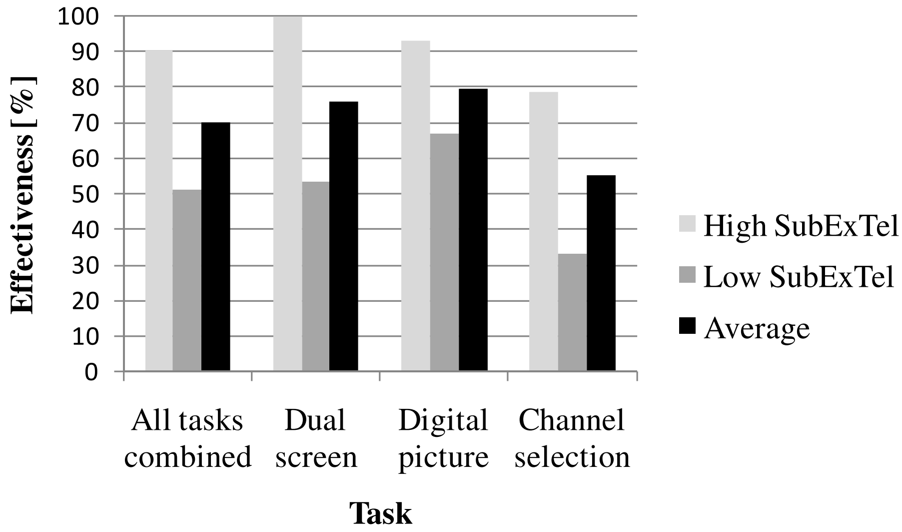


Figure 5.9 Overview of effectiveness measures when differentiating on subjective expertise of TVs

Next, a separate pair-wise Mann-Whitney U test showed a significant overall effect of subjective expertise of TVs on task completion ($p < 0.001$), which supports hypothesis H_1. For all the tasks a higher percentage of subjects of the high subjective expertise group were able to complete the task than subjects of the low subjective expertise group. Specific subjective expertise differences for each separate task are discussed in the following sections.

Effect of subjective expertise and task complexity on satisfaction

An overview of differences in satisfaction between the tasks and the subjective expertise groups is shown in Figure 5.10. One-way Kruskal-Wallis tests showed no significant effect of task complexity on satisfaction (χ^2 (2) = 2.887, $p < 0.3$) and a separate pair-wise Mann-Whitney U test showed only a marginal significant effect of subjective expertise on satisfaction ($p < 0.1$). For almost all the tasks and separate groups the satisfaction scores are above the average of the scale (i.e. above four), which indicates a modest level of satisfaction with the usability of the LCD TV. Since no significant differences could be observed, no further analyses for this variable are discussed.

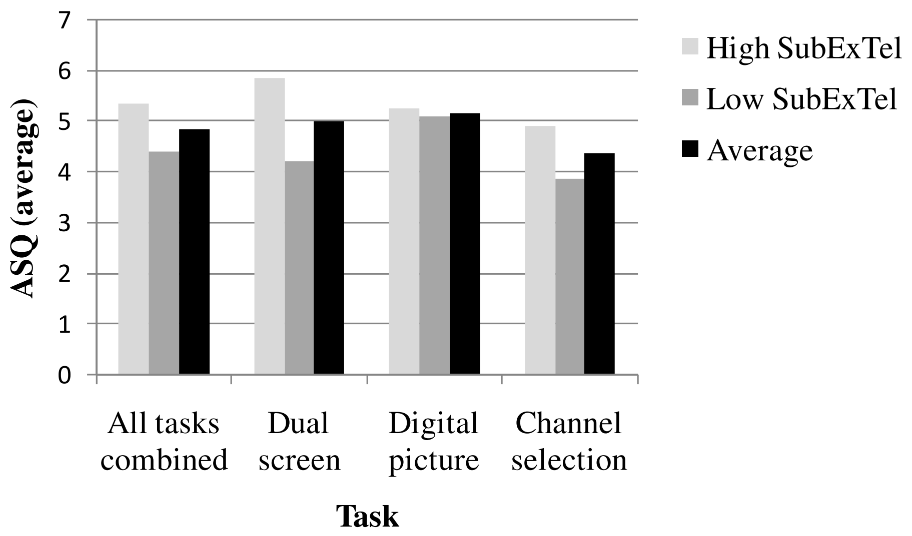


Figure 5.10 Overview of average level of satisfaction when differentiating on subjective expertise of TVs

Overall effect of objective familiarity
To test the final hypothesis on expected differences between objective familiarity and subjective expertise as predictors of product usage behavior, separate analyses of the effect of objective familiarity (on both the core and supplemental domain) on usage behavior are performed. The MANOVA test of a model with task complexity and objective familiarity of TVs (two groups based on a split on the mean value across all the test participants) as an independent variable did not show an overall significant effect of objective familiarity of TVs on the efficiency measures ($F (3, 66) = 0.880$, $p < 0.5$). A separate pair-wise Mann-Whitney U test indicated that there was only a marginally significant effect of objective familiarity of TVs on effectiveness ($p < 0.1$). Finally, no significant differences between separate tasks could be observed.

However, a second MANOVA test of a model with objective familiarity of computers and task complexity as independent variables did show a significant overall effect of objective familiarity with computers on the efficiency measures ($F (3, 66) = 2.960$, $p < 0.05$) and also of task complexity ($F (6, 134) = 4.268$, $p < 0.005$). Subsequent univariate tests indicated that differentiation of objective familiarity of computers only had a significant effect on time ($F (1, 74) = 5.692$, $p < 0.05$) and not on the number of steps ($F (1, 74) = 0.207$, $p < 0.7$) and number of detour steps ($F (1, 74) = 0.000$, $p < 1.0$). Furthermore, a separate pair-wise Mann-Whitney U test showed that there was no significant overall effect of objective familiarity of computers on effectiveness ($p < 0.2$) and also no significant differences between separate tasks could be observed.

Based on these results can be concluded that H_3 can be accepted because the results for subjective expertise demonstrated a stronger impact on product usage behavior than the results for objective familiarity. Furthermore, H_4 needs to be rejected because differences in objective familiarity of TVs have no impact on product usage behavior while the impact of objective familiarity of computers was significant.

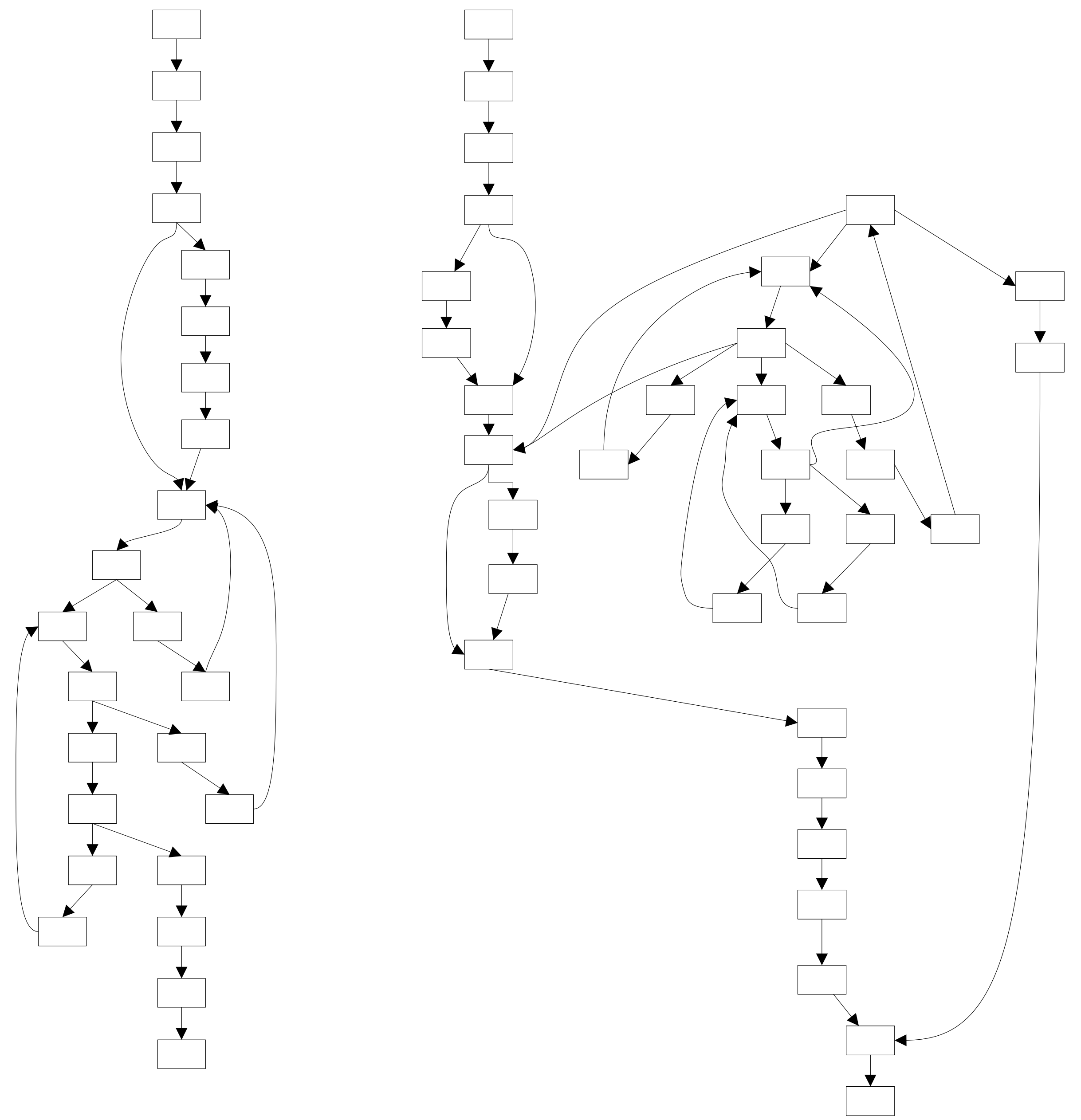
5.3.3 Dual screen task

In this section, the specific results and observations for the dual screen task are discussed. An overview of the descriptive statistics of the dependent variables for this task is shown in Table 5.5. Results of separate pair-wise Mann-Whitney U tests showed significant differences between high and low levels of subjective expertise for effectiveness ($p < 0.05$) and time ($p < 0.05$), and marginally significant differences for steps ($p < 0.1$), levelup ($p < 0.1$) and satisfaction ($p < 0.1$). As was shown in Figure 5.9, in terms of effectiveness all the participants with high subjective expertise were able to complete the task while for the low subjective expertise group only just over half the participants completed the task. For subjective expertise on computers similar significant differences were observed while for objective familiarity no significant differences were observed.

Table 5.5 Descriptive statistics of the dependent variables for the dual screen task

	High subjective expertise on televisions			**Low subjective expertise on televisions**		
	Mean	**S.D.**	**Range**	**Mean**	**S.D.**	**Range**
Time [seconds]	154.07	153.127	51 – 601	355.08	242.928	51 – 715
Steps [total number]	6.36	4.308	4 – 18	12.25	11.663	4 – 45
Levelup [total number]	0.71	1.490	0 – 4	2.58	4.078	0 – 14
Satisfaction [scale average]	5.86	1.300	3 – 7	4.22	2.285	1 – 7

The results are further supported by the mined heuristic models and the performance sequence diagram analysis of ProM. In Figure 5.11 on the following page, the mined models for this task (based on the event logs) for respectively the high and low subjective expertise groups are shown. Please note that the purpose of showing these models is to visualize differences between the subjective expertise groups in terms of how many events took place, in which sequence events took place and which loops (return to previous state) occurred. When both models are visually compared can be seen that more different events and more inefficient loops were logged for the low subjective expertise group. As a main cause was observed that many participants from the low subjective expertise group started the task by looking for the dual screen functionality in the TV menu while a separate remote control button could be used to access this function directly.



High subjective expertise Low subjective expertise

Figure 5.11 Mined heuristic models of dual screen task for the high subjective expertise group (left) and the low subjective expertise group (right)

Although the functionality could be accessed in a lower level in the TV menu, the functionality could not be found and/or the manual was unclear on where to locate the function. Furthermore, at the moment the functionality was found, many participants from the low subjective expertise group had difficulty finding out how to switch between the left and right part of the dual screen to change both channels.

The results are also reflected in the performance sequence diagram analysis shown in Figure 5.12 on the next page. The X-axis of this diagram shows the states or functions which were visited by the participants and the Y-axis shows the time spent in each state (please note that only sequences from completed tasks could be included in the analysis). The diagram groups similar patterns of sequential steps together and thus only patterns that occurred more than once have been included in the diagram. An overview of the statistics for the three patterns found is shown in Table 5.6.

Table 5.6 Performance sequence diagram statistics for the dual screen task

	Pattern 0	Pattern 1	Pattern 2
Average throughput time (s)	164.07	69.32	126.86
Minimum throughput time (s)	54.20	52.33	53.24
Maximum throughput time (s)	438.43	105.36	200.48
S.D. throughput time (s)	150.18	22.56	104.11
Frequency	7	5	2

In the diagram can be seen that both in “pattern 0” and in “pattern 2” (with most participants of the low subjective expertise group) a large amount of time was spent in “start task” and “dual screen mode” due to a lack of understanding of the functionality and looking for more information on how to use this function in the manual or on the remote control. For the participants with a low level of subjective expertise both the design of the UI of the LCD TV and the information in the manual did not provide enough support to complete the task. “Pattern 1” contains participants from the high subjective expertise group and displays a reasonable efficient pattern. Consequently, all the results combined show that the dual screen task participants from the low subjective expertise group were significantly less effective and efficient on all the measurements compared to the participants from the high subjective expertise group.

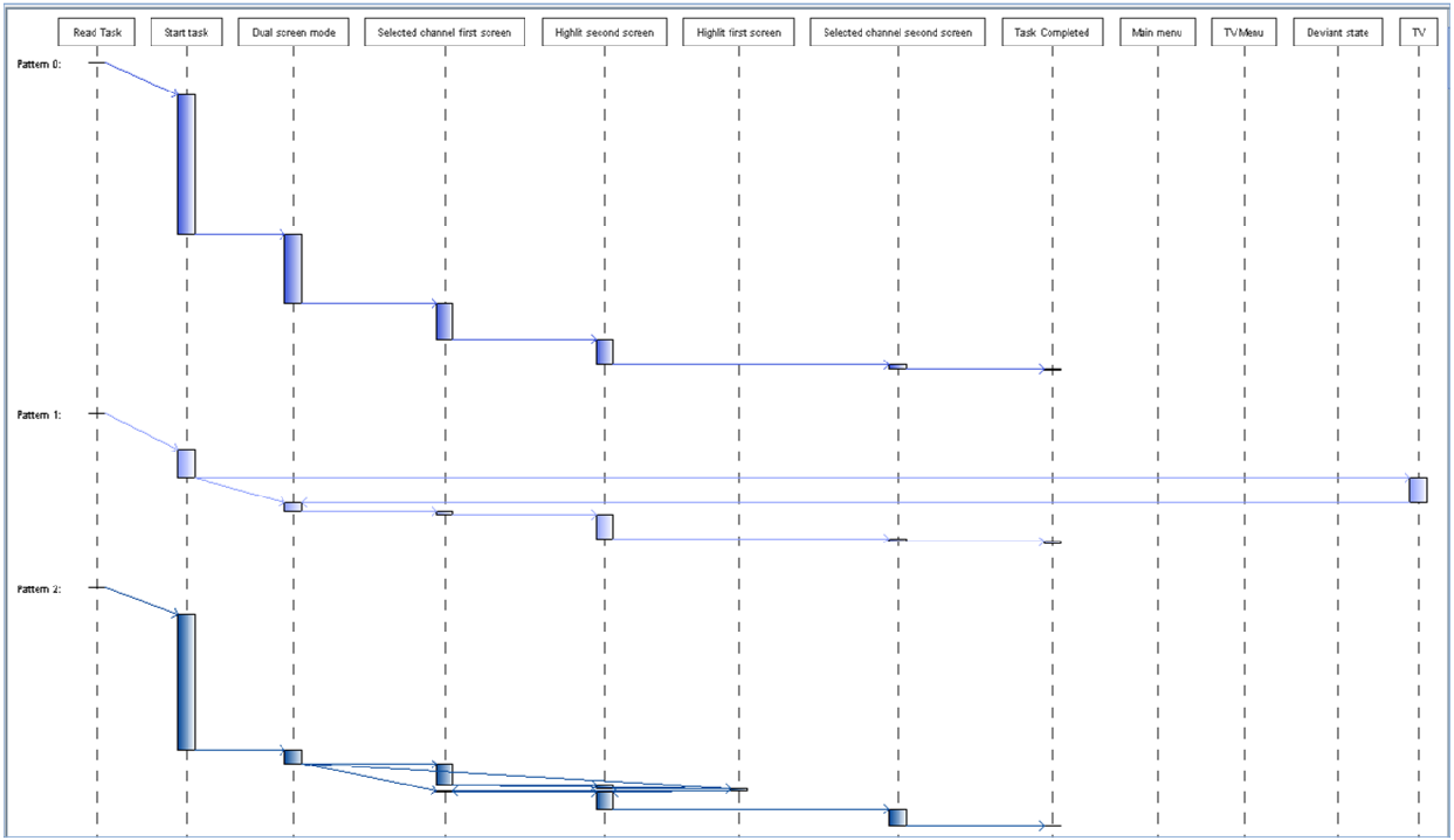


Figure 5.12 Performance sequence diagram for dual screen task

5.3.4 Digital picture task

In contrast to the dual screen task, for the digital picture task separate pair-wise Mann-Whitney U tests did not show any significant differences between high and low subjective expertise of TVs for all of the dependent variables. An overview of the descriptive statistics for the dependent variables based on this differentiation is shown in Table 5.7. However, additional analyses for the effect of differentiation on subjective expertise of computers and objective familiarity on both knowledge domains did result in a significant difference on time ($p < 0.05$) when differentiating on subjective expertise of computers and a marginally significant different on time ($p < 0.1$) when differentiating on objective familiarity of computers. In other words, for this task, which originates from the computer domain, the knowledge on this supplemental domain has a stronger (albeit small) effect on efficiency than knowledge on the core domain.

Table 5.7 Descriptive statistics of the dependent variables for the digital picture task

	High subjective expertise on televisions			Low subjective expertise on televisions		
	Mean	**S.D.**	**Range**	**Mean**	**S.D.**	**Range**
Time [seconds]	206.14	144.014	65 – 476	288.47	145.899	83 – 561
Steps [total number]	11.64	7.344	4 – 34	9.00	5.490	4 – 21
Levelup [total number]	1.71	1.204	0 – 4	1.07	0.961	0 – 3
Satisfaction [scale average]	5.26	1.508	1 – 7	5.09	1.740	3 – 7

The mined models using the heuristic miner and the results of the performance sequence diagram analysis for this task are shown in Appendix 5.7 and Appendix 5.8 respectively. Similar to the dual screen task, for this task, one of the identified reoccurring performance sequences, for mostly the low subjective expertise group, is in which the task starts with a long time spent in the "start task" state where participants read the manual to try to understand how to perform this task. However, in contrast to the dual screen task, the mined models did not show a larger number of loops and deviant states for the low subjective expertise group. Two possible explanations could be found for this (unexpected) lack of significant differences:

- The design of the experiment was such that a USB stick was already provided at the start of the task. This could have affected the efficiency measurements positively for the low subjective expertise group because the first step ("identifying" a USB stick) was already "completed" before the start of the task.
- The multimedia browser in which the externally connected multimedia devices were displayed on the UI of the LCD TV showed two different USB sticks and did not

highlight which port was used by the USB stick with the photos. Several participants from the high subjective expertise group first investigated the content of both to see which one should be used to complete the task while participants from the low subjective expertise group either randomly selected a USB port or followed the manual step by step.

5.3.5 Switch channel task

The descriptive statistics of the dependent variables for the switch channel task are shown in Table 5.8. Results of separate pair-wise Mann-Whitney U tests showed significant differences between high and low levels of subjective expertise for effectiveness ($p < 0.05$) and time ($p < 0.05$), but no significant differences for steps ($p < 0.3$), levelup ($p < 0.3$) and satisfaction ($p < 0.2$). As was shown in Figure 5.9, in terms of effectiveness 78.8% of the participants with high subjective expertise were able to complete the task while for the low subjective expertise group only one third of the participants completed the task. For subjective expertise on computers similar significant differences were observed while for objective familiarity (of both the TV and computer domain) no significant differences were observed.

Table 5.8 Descriptive statistics of the dependent variables for the switch channel task

	High subjective expertise on televisions			**Low subjective expertise on televisions**		
	Mean	**S.D.**	**Range**	**Mean**	**S.D.**	**Range**
Time [seconds]	237.00	153.127	51 – 601	533.13	284.258	163 – 903
Steps [total number]	20.18	13.220	7 – 49	28.75	18.109	7 – 68
Levelup [total number]	3.82	3.970	0 – 12	6.75	4.862	0 – 16
Satisfaction [scale average]	5.61	1.223	3 – 7	4.64	2.015	3 – 7

The mined models using the heuristic miner and the results of the performance sequence diagram analysis for this task are shown in Appendix 5.7 and Appendix 5.8 respectively. First of all, for this most complex task of the three tasks, the results of the performance sequence analysis only showed one sequence with multiple occurrences ($n = 3$) indicating that for both subjective expertise groups many different sequences of steps were used which could not be grouped together. Although both mined models show loops and not many other differences between the models could be observed, further analysis of the results indicates that participants from the low subjective expertise group:

- Had more difficulty finding the functionality to switch channels (due to the many layers in the menu) and more often started to reprogram the channel manually (which could have resulted in the same requested end result for this task but was far more complex to use).

- Had more difficulty to use the remote control to select and subsequently switch channels.

Finally, a usability problem emerging from the analysis is that most of the participants experienced some difficulty in understanding the TV menu for this function because they had difficulty to understand the Dutch translation of the word for "rearrange channels" (i.e. translated as "zenders herschikken") in the menu.

5.4 Conclusion and discussion

This section concludes this chapter and discusses the results of the study and its implications. First, section 5.4.1 gives an overview of the results and discusses the hypotheses. Subsequently, section 5.4.2 discusses how the results impact the value of selecting consumers for product tests based on their level of product familiarity. In section 5.4.3 several implications of consumer knowledge differences for product design are discussed. Finally, in section 5.4.4 limitations of the study and implications for further research are discussed.

5.4.1 Overview of the results

This study investigated how both subjective expertise and objective familiarity differences on the core and supplemental knowledge domain of a multimedia LCD TV affect product usage behavior for tasks with varying level of complexity. These differences were evaluated in a laboratory experiment with 29 participants recruited from the respondents of the survey discussed in Chapter 4.

The results demonstrated that overall the design of the experiment was successful: task complexity differences were reflected in differences in effectiveness and efficiency and separate analyses of differences in intention-to-use showed no significant effect of this potentially confounding variable. The potentially moderating effect of age was also evaluated but did not improve the explanatory power of the statistical analysis. An overview of the hypotheses tested in this study is shown in Table 5.9. From this table can be seen that most of the defined hypotheses on the effect of subjective expertise on effectiveness and efficiency are confirmed. Overall, the positive effect of subjective expertise on the ability to complete tasks is supported. Similarly, participants with a high level of subjective expertise were overall found to complete tasks in less time, with a fewer number of detour steps and with usage patterns which conformed more to the designed usage model in comparison with participants with a low level of subjective expertise. Hypothesis H_{2b} concerning differences in the number of steps needed to complete a task was not supported. Furthermore, in contrast to Ziefle (2002) no significant interaction effects between task complexity and consumer knowledge were observed. In other words, the effect of consumer knowledge was equal across different levels of task complexity.

Table 5.9 Overview of results of hypotheses testing in Chapter 5

Hypothesis	Result
H_1: Consumers with higher levels of knowledge, measured as follows: a) Subjective expertise b) Objective familiarity of TVs c) Objective familiarity of computers perform tasks more effectively than do consumers with lower levels of the same measure of knowledge	a) Accepted b) Rejected c) Rejected
H_{2a}: Consumers with higher levels of knowledge, measured as follows: a) Subjective expertise b) Objective familiarity of TVs c) Objective familiarity of computers need less time to perform tasks than do consumers with lower levels of the same measure of knowledge	a) Accepted b) Rejected c) Accepted
H_{2b}: Consumers with higher levels of knowledge, measured as follows: a) Subjective expertise b) Objective familiarity of TVs c) Objective familiarity of computers need less steps to perform tasks than do consumers with lower levels of the same measure of knowledge	a) Rejected b) Rejected c) Rejected
H_{2c}: Consumers with higher levels of knowledge, measured as follows: a) Subjective expertise b) Objective familiarity of TVs c) Objective familiarity of computers make less detour steps than do consumers with lower levels of the same measure of knowledge	a) Accepted b) Rejected c) Accepted
H_{2d}: Consumers with higher levels of knowledge, measured as follows: a) Subjective expertise b) Objective familiarity of TVs c) Objective familiarity of computers conform more to the ideal usage pattern than do consumers with lower levels of the same measure of knowledge	a) Accepted b) Not tested c) Not tested
H_3: Differences in subjectively measured expertise have a stronger effect on product usage behavior than differences in objectively measured familiarity	Accepted
H_4: Differences in objective familiarity of televisions have a stronger effect on product usage behavior than differences in objective familiarity of computers	Rejected

Further analyses of the logged events using ProM showed that for the dual screen and switch channel tasks participants from the low subjective expertise group experienced more and different interaction problems than participants from the high subjective expertise group.

However, this effect was not fully observed for the digital picture task. Two possible explanations for this inconsistency with the results of the other two tasks were the experimental design of the task (providing the USB stick is already one completed step towards completing the task without necessarily understanding what a USB stick is and how to use it) and the design of the multimedia browser interface. Although the UI was designed that it automatically selected the right USB port, several participants of the high subjective expertise group checked the content of both ports before proceeding to the stored pictures.

Next, the results of this study show that the models incorporating subjective expertise, in general, have a stronger and more significant impact on usage behavior than models with objective familiarity. This is in line with Brucks (1985) and Cordell (1997) who argue in a different research context that, although familiarity is often used as a proxy measure for consumer knowledge, differences in expertise stronger relate to differences in factual knowledge and subsequent consumer behavior than differences in familiarity.

Finally, the results of additional analyses with objective familiarity as independent variable demonstrate that, in contrast to the hypothesis, objective familiarity of computers has a significant overall effect on the efficiency measurements and objective familiarity of TVs does not have a significant effect. Similar to the discussion on reasons for the very low correlation between objective familiarity and subjective expertise of TVs in section 4.4, one could argue that usage experience of TVs mostly relates to a relatively "passive" form of interaction while usage experience of computers in general can encompass many different and also more active forms of interactions. Consequently, although for subjective expertise measurements this was not useful due to strong correlation, the use of core and supplemental knowledge domains can help to explain differences in product usage behavior of a CE product which merges functionalities from historically different product domains (although the effect is small).

5.4.2 Selecting consumers for product tests based on a differentiation on consumer knowledge

This study demonstrates that consumer knowledge differences significantly impact how consumers respond to complex tasks in CE and as such this construct can be used as a differentiator for selecting consumers for product tests. Currently, in usability literature and practice, differentiating on usage experience (i.e. experts vs. novice) is common practice when testing products for usability problems (e.g. Nielsen (1993)). However, the results of this study showed that, for LCD TVs, subjective expertise has a stronger effect on usage behavior than objective familiarity which is similar to usage experience. Moreover, in the high subjective expertise group there were two participants who rated themselves as having a high level of subjective expertise but in terms of performance could be rated as a product novice. Although the limited sample size did not allow for further investigation, this so-called miscalibration of knowledge (i.e. having the perception of knowing a lot about how a product functions but factually knowing less) (Alba & Hutchinson, 2000; Carlson et al., 2009) could have a more significant effect on product usage behavior than differences in usage experience.

Consequently, taking subjective expertise as a differentiator of consumers for product tests into account could potentially add value, especially for complex technological products which rapidly change from a product technological point of view but not necessarily from a consumer point of view. On top of that, subjective expertise measurements are easy to apply in practice and do not require complex, tailored measurements such as required for objective expertise.

5.4.3 Considering consumer knowledge differences in product design

Besides having implications for selecting consumers for product tests, the results of this study also confirm and further enrich earlier research which shows that consumer knowledge differences have to be taken into account during product design (e.g. Nielsen, 1993; Ziefle, 2002). The results of the study demonstrated that participants in the low subjective expertise group not only encountered more and different interaction problems but also used different strategies to complete a task than the participants in the high subjective expertise group. On average they took more time to look for information on how to use a certain function and had more difficulty to understand and navigate through the TV menu (e.g. not understanding how to use the remote control to navigate through the menu or to select a certain function). Since CE such as LCD TVs are developed for mass consumer markets, taking consumer knowledge differences into account can help to tailor both manual (e.g. see Novick and Ward (2006)) and UI design (e.g. see Belkin (2000)) to account for consumer diversity.

5.4.4 Study limitations and implications for further research

The study presented in this chapter has given more insight into how consumer knowledge differences affect product usage behavior when consumers are confronted with the inherent complexity of CE. Nevertheless, this study had several limitations. First of all, only familiarity and subjective expertise measurements were used. Although theory suggested that subjective measurements of expertise are adequate proxy measures of objective expertise for luxury and durable goods (Carlson et al., 2009), for future studies it would give more information to also include objective measurements of expertise. Furthermore, only a limited sample size could be achieved which also resulted in relatively small effect sizes. Since consumer knowledge must be studied in combination with potentially moderating factors and because of the quasi-experimental research methodology, larger sample sizes for future studies are required.

These limitations will be addressed in Chapters 6 and 7 in which the effect of consumer knowledge on failure attribution is investigated. In both studies simulated product failures are taken into account and besides familiarity and subjective expertise measurements also an objective expertise measurement of LCD TVs is used.

6 Evaluating the effect of consumer knowledge and failure origin on failure attribution[17]

After having investigated the effect of consumer knowledge on product usage behavior, this chapter and the following chapter investigate the effect of consumer knowledge on failure attribution. Both chapters explore this effect for different failure characteristics using different research methodologies. This chapter specifically investigates how consumer knowledge differences and the how physical cause of a failure affect failure attribution using an Internet-based experiment with implemented videos of failure scenarios.

This chapter is organized as follows. Section 6.1 discusses the research variables used in this study. This section concludes with a conceptual framework of this study and hypotheses are formulated. Subsequently, in section 6.2 the design of the Internet-based experiment to test these hypotheses is discussed. Section 6.3 reports on the results of this experiment, assessing the effect of both the consumer knowledge constructs and failure origin on failure attribution. Finally, this chapter concludes with a discussion of the results and limitations of this study in section 6.4.

6.1 Conceptual framework and hypotheses

This section discusses the set-up of the conceptual framework and the selection of the research variables to investigate the effect of consumer knowledge and failure cause on failure attribution. Section 6.1.1 discusses the overall goals of this study with respect to attribution of failures in DTV systems. Based on the goals of this study and earlier research, section 6.1.2 discusses the selection of consumer knowledge measurements. Subsequently, similar to the previous studies, section 6.1.3 discusses the selection of control and moderating variables. Finally, in section 6.1.4 the conceptual research framework and hypotheses are presented.

6.1.1 Attribution of failures in DTV systems

As discussed in Chapter 1, product development faults are only important when they are triggered during product use, attributed as a failure and subsequently result in consumer

[17] Part of the material presented in this chapter is published in: "Keijzers, J., Den Ouden, P.H. & Lu, Y. (2009). Understanding consumer perception of technological product failures: An attributional approach. *In Proceedings of the 27th International Conference Extended Abstracts on Human Factors in Computing Systems*, (pp. 4057–4062). New York: ACM".

dissatisfaction. While the previous chapters focused on the effect of consumer knowledge on the product usage behavior, this and the following chapter focus on the effect of consumer knowledge on the second aspect: failure attribution. For example, research by Ceaparu, Lazar, Bessiere, Robinson and Schneiderman (2004) has shown that novice and even expert consumers can wrongly interpret a product's behavior, product (error) feedback messages and even the manual, which could lead to an ineffective problem solving strategy or even more consumer frustration. Since product designers have difficulty predicting this consumer behavior with respect to product failures, the goal of both failure attribution studies discussed in Chapter 6 and 7 is to gain better insight into how different consumer groups perceive (potential) product failures. This insight can ultimately be used to support design decisions in the future product development processes. The first step is to investigate the differences in the perceived failure causes between consumers with different levels of consumer knowledge and subsequently comparing this failure attribution with the real physical cause of the failure. In other words, consumer knowledge differences are expected to be a main determinant of how consumers attribute problems that might not be in accordance with the real physical cause as determined by product experts.

In Chapter 3 the (hypothesized) relation between consumer knowledge and failure attribution has been discussed without explicitly referring to different types of product failures. Furthermore, in Chapter 1 and 2 it was discussed that CPFs can have multiple causes: product development faults, the environment, the consumer or a combination of these. Similarly, each of these different CPFs can be perceived by the consumer to be caused by the product, the environment, the consumer or a combination of these. However, for practical reasons it is neither possible nor desirable to investigate the consumer's attribution of each of these antecedents of CPFs; especially since it is important, for reasons of experimental validity (Stangor, 1998, p. 158), to control for extraneous variables such as the characteristics of the failure (De Visser, 2008, p. 67) and the use conditions. For example, research by Laufer, Gillespie, McBride & Gonzalez (2005) has shown that severity of a failure can influence the extremity of attribution.

In the context of the TRADER project there is specific interest in the consumer's perception of potential software failures in DTV systems and in the consumer's perception of potential "side-effects" of software failure recovery mechanisms that might lead to other CPFs. Recovery mechanisms are used to prevent the occurrence of failures but it might be that, in the perception of the consumer, the consequences of the recovery in terms of observable product functioning (e.g. a short picture freeze or a temporary degraded picture quality) are worse than the failure that the recovery mechanism is trying to prevent. In this context, only failures that could be physically caused by either the (product usage) environment or the product itself but not by consumer or consumer-product interaction were considered. The latter would be practically not feasible to study in a laboratory environment and would not allow for sufficient control of extraneous variables.

Consequently, for this first study on the effect of consumer knowledge on failure attribution it was decided to specifically focus on two failures with a distinctly different physical cause; i.e. a physical cause internal to the TV due to a software fault and a physical cause external to the TV (e.g. due to a fault in a DVD or the cable signal). A choice is made to further limit the scope of this study and the following study discussed in Chapter 7 to failures in picture quality of an LCD TV. A survey by De Visser (2008, p. 79) has shown that watching a desired program is the most important function of a multimedia LCD TV from a consumer point of view and hence failures in this function are expected to have the strongest influence on attribution processing for this product domain. In this context, the goal of this study is to investigate how and to what extent consumer knowledge affects differences in attribution of both failures in TV picture quality and subsequently to investigate whether the hypothesized reasons for the failure by the consumer match the determined physical cause of the failure by the DTV system experts. The selection and design of the failure scenarios will be further discussed in section 6.2.3.

As discussed in section 3.3.2, although attribution locus (i.e. do consumers perceive the cause of the failure to be internal or external to the TV) is of interest for this research, for reasons of external validity the other attribution dimensions also need to be taken into account when evaluating attribution differences (Folkes, 1988).

6.1.2 Selection of consumer knowledge measurements

In the studies discussed in Chapters 4 and 5 it was investigated how consumers can be differentiated on *subjective expertise*, *subjective familiarity* and *objective familiarity* of the core and supplemental knowledge domain of LCD TVs and subsequently how differences on these consumer knowledge measurements affect product usage behavior. Objective expertise measurements were not taken into account in previous studies because theoretically it was argued that subjective expertise would have a stronger effect on product usage behavior and because it would have required too much time to develop such a measurement already in the explorative stage of this research. For the study discussed in this chapter all consumer knowledge measurements were taken into account (i.e. objective and subjective measurements of expertise and familiarity) for two main reasons:

- Previous research has shown that objective expertise has a significant effect on information processing (which relates to processing of attribution of failures) and that this effect is significantly different from subjective expertise (Brucks, 1985; Sujan, 1985). Furthermore, a study by Somasundaram (1993) has shown that higher levels of objective expertise result in a greater ability to generate plausible causes to explain a failure in photograph processing.
- Taking into account objective expertise differences enables a more complete evaluation of the first sub research question on how consumers can be differentiated on consumer knowledge of CE.

Furthermore, in contrast to the previous study the choice is made to evaluate only the effect of consumer knowledge measurements of the LCD TV domain and no longer differentiate between core and supplemental knowledge domains. First of all, the results of the previous studies demonstrated that there is a high correlation between subjective expertise measurements of the TV and computer domain and also a similar effect on product usage behavior was observed. Secondly, since an objective expertise measurement is idiosyncratic with the product class and should be tailored to the research variables of interest in a study (Alba & Hutchinson, 1987; Cordell, 1997), an objective expertise measurement of LCD TVs can include knowledge on the core functionalities of a TV as well as on software content and the technical functioning of LCD TVs.

6.1.3 Selection of control variables and moderating variables

As similarly stated in section 4.1.4, to be able to fully understand the relationship between consumer knowledge and failure attribution several moderating and control variables need to be taken into account (Alba and Hutchinson, 1987; Carlson et al., 2009; Folkes, 1988; Silvera & Laufer, 2005).

First of all, the results of the consumer knowledge survey discussed in Chapter 4 showed that age difference should be taken into account because this demographic variable significantly correlated with subjective expertise and familiarity. Furthermore, research has shown that age might influence cognitive performance (in this context the ability to reason about the perceived failure cause) and biased processing of attribution (Laufer, Silvera & Meyer, 2005).

Secondly, as discussed in section 3.3.2, there are three types of antecedents which influence failure attributions (Kelley & Michaela, 1980; Folkes, 1988). Since in this dissertation the main focus is on the effect of information as an antecedent (i.e. knowledge regarding a particular product and its potential failures), the effect of the following two antecedents needs to be either controlled for or needs to be taken into account in the analysis (Kelley & Michaela, 1980; Folkes, 1988):

- Product involvement (relates to motivation to think about causal relations): in the context of causal attribution search, consumers with a higher level of involvement are more likely to think about the causes of a product failure (Somasundaram, 1993).
- Product expectations (relates to prior beliefs): erroneous or extreme expectations or hypotheses of a product’s performance might influence the attribution of failures related to that product.

Furthermore, although failure cause is the main failure characteristic of interest for this study and although other failure characteristics need to be controlled for in the experimental design, it is important to take the perceived severity of a failure into account. On the one hand, although previous research results are inconclusive (Sujan, 1985; Somasundaram, 1993), consumer knowledge could affect the extremity of a failure evaluation and on the other hand perceived severity can in turn affect failure attribution (Laufer, Gillespie et al., 2005). Therefore, measurements of both failure impact (i.e. perceived degree of loss of functionality (De Visser, 2008, p. 68) and perceived picture quality are included in the conceptual framework to investigate these potentially confounding effects.

A final control variable which needs to be taken into account is failure experience. Because on the one hand a quasi-experimental approach is used and because on the other hand failure experience is expected to covary with consumer knowledge, it is not possible to control for prior experience with a TV failure. Although prior experience (in a sense this can be referred to as "failure familiarity") does not necessarily result in a more correct attribution of a failure, this potential effect needs to be taken into account during the analysis since it might lessen or strengthen the effect of consumer knowledge on attribution.

6.1.4 Conclusion

Summarizing, the study discussed in this chapter serves two different goals. First, to set-up and validate an objective expertise measurement of LCD TVs and to investigate the differentiation of consumers into segments based upon this measurement. Secondly, to investigate how consumer knowledge differences affect failure attribution of failures with a different physical cause (i.e. internal or external to the product) regarding the picture quality of an LCD TV.

For the second goal, the research variables and its hypothesized relations discussed in this section highlighted in the overall research model as shown in Figure 6.1. This research model will be used (together with the results of the study discussed in Chapter 7) in the study to answer the third research sub question defined in section 3.4. The results of the literature review on consumer knowledge and failure attribution discussed in section 3.2 and 3.3 indicated that consumers are expected to attribute product failures to explanations that are consistent with the consumer's existing knowledge (Oliver, 1996). Since research has shown that higher levels of expertise, among other things, are manifested in more refined cognitive structures of a product, an increased ability to distinguish relevant information and an increased ability to generate more elaborate explanations (Alba & Hutchinson, 1987), it is generally expected that higher levels of consumer knowledge result in more correct[18] but also more refined and/or elaborate attributions independent of the type of failure. Consequently, it is hypothesized that:

Hypothesis H_1: Consumers with higher levels of knowledge, measured as follows:

a) Objective expertise
b) Subjective expertise
c) Objective familiarity
d) Subjective familiarity

[18] In this context "correct" implies more in accordance with the physical cause of the failure.

attribute product failures caused by product internal factors stronger to internal causes than do consumers with lower levels of the same measure of knowledge.

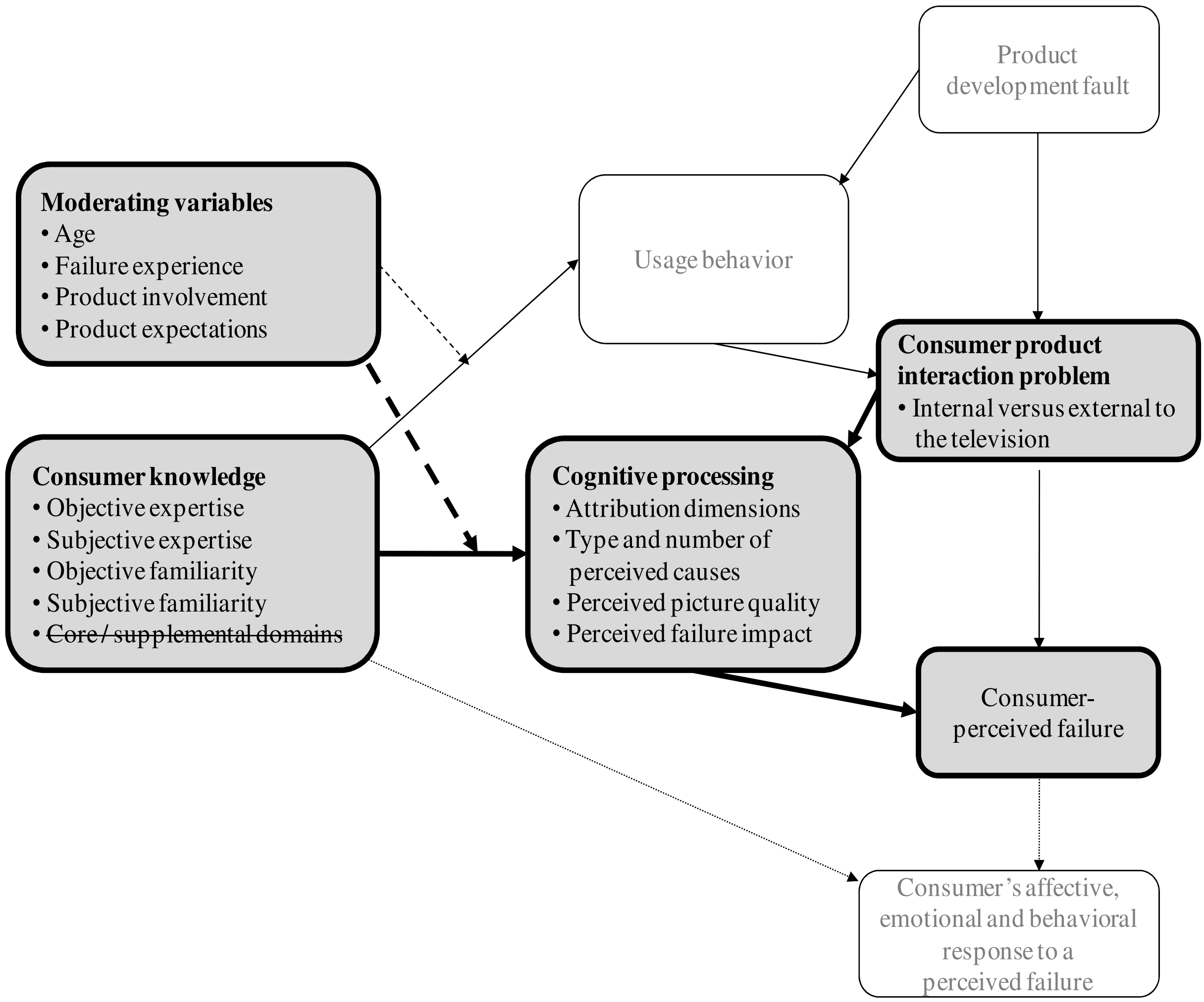


Figure 6.1 Conceptual framework to investigate the effects of consumer knowledge and failure origin on failure attribution

Hypothesis H_2: Consumers with higher levels of knowledge, measured as follows:

e) Objective expertise
f) Subjective expertise
g) Objective familiarity
h) Subjective familiarity

attribute product failures caused by product external factors stronger to external causes than do consumers with lower levels of the same measure of knowledge.

Hypothesis H_3: Consumers with higher levels of knowledge, measured as follows:

a) Objective expertise
b) Subjective expertise
c) Objective familiarity

d) Subjective familiarity

attribute product failures to more different causes than do consumers with lower levels the same measure of knowledge.

Besides hypotheses on the overall effect of consumer knowledge on failure attribution, it is also interesting to investigate which consumer knowledge measurement most strongly relates to differences in failure attribution. The results of the experiment discussed in Chapter 5 and previous research by Brucks (1985) and Cordell (1997) have shown that the expertise based component of consumer knowledge is more directly related to behavior and product evaluation than familiarity. Furthermore, because objective expertise is regarded as the most reliable measure of what people *actually* know (Brucks, 1985; Cordell, 1997) and reflects qualitative aspects of expertise (Arning & Ziefle, 2009), it is expected to be stronger related to differences in failure attribution than subjective expertise. This resulted in the following hypothesis:

Hypothesis H_4: Differences in objectively measured expertise stronger relate to differences in failure attribution than differences in familiarity and subjectively measured expertise.

Finally, since previous research is inconclusive on the direction of the effect of consumer knowledge on perceived severity of a failure and extremity of beliefs related to product evaluation (Somasundaram, 1993; Sujan, 1985), the following two hypotheses were formulated:

Hypothesis H_5: Consumers with higher levels of knowledge , measured as follows:

a) Objective expertise
b) Subjective expertise
c) Objective familiarity
d) Subjective familiarity

rate perceived picture quality differently from consumers with lower levels of the same measure of knowledge.

Hypothesis H_6: Consumers with higher levels of knowledge, measured as follows:

a) Objective expertise
b) Subjective expertise
c) Objective familiarity
d) Subjective familiarity

rate perceived failure impact differently from consumers with lower levels of the same measure of knowledge.

6.2 Method

To test the hypotheses formulated in the previous section, a 2 (high versus low levels of consumer knowledge) x 2 (internally versus externally caused product failure) between-subjects Internet-based experiment was designed in which participants were asked to evaluate a scenario showing a product failure related to picture quality of an LCD TV. This section describes the selection of the research methodology and set-up of this experiment.

6.2.1 Research methodology

There are several commonly used research methodologies to investigate consumers' reactions to product or service failures. Based on the literature reviews by Ma (2007) and Lancellotti (2004) and the discussion on failure attribution research by Weiner (2000), three distinctly different methodologies were identified:

- Retrospective survey: Recollection of the respondents' perception of the antecedents and/or consequences of a product or service failure that was experienced in the recent past.
- Scenario-driven experiment: Role-playing experiment in which participants are asked to read a short description of a failure scenario and respond to questions regarding their perception, attribution etc. Alternative, but similar methods include the use of online scenarios and simulations.
- Laboratory experiment: Experiment in which the participants are asked to use a certain product or service (usually embedded in tasks which do not explicitly focus on failures). During the task an implemented failure occurs to which the participants are asked to respond.

The retrospective survey was not suitable for this study to be able to evaluate the effect of failure characteristics on attribution, a controlled failure scenario with a known and fixed failure cause was required. Based on a comparison of the methodological and practical advantages and disadvantages of the scenario-driven versus the laboratory experiment, a choice was made to use a scenario-driven Internet-based experiment with video clips of simulated failures. This method allowed for easier manipulation of the failure scenarios, greater control of extraneous variables and easier reach of a relatively large sample size (required because of the use of the quasi-experimental research methodology) compared to the laboratory experiment. Although a scenario-driven Internet-based experiment has less ecological validity because it does not involve a real-life failure experience, it does enable the investigation of failure attribution for a large and diverse group of participants from various backgrounds and a potentially large spread on consumer knowledge. Furthermore, Internet-based experiments also suffer from both practical and methodological drawbacks for which specific attention in the design of the experiment and extensive pre-tests are needed (Manfreda & Vehovar, 2008; Reips, 2002a; Reips, 2002b). This will be further discussed in the following sections.

6.2.2 Research variables: Consumer knowledge measurements

Since the four different consumer knowledge measurements and the failure attribution measurements and scenarios needed to be included in one single Internet-based questionnaire, it was important to keep the measurements of the research variables as short as possible to achieve a higher response rate (Dillman, 2000, p. 305). For the measurements of subjective expertise, subjective familiarity and objective familiarity the insights gained from the survey discussed in Chapter 4 were used to adjust and/or shorten the measurement scales.

For the subjective expertise measurement the two different items which scored highest on the adjusted subjective expertise scale from Flynn and Goldsmith (1999) were used. Furthermore, for the subjective familiarity measurement only adjusted subjective TV usage items were used since the results of the survey discussed in Chapter 4 showed that information search items did not score on one construct. Also, for this construct the number of items was reduced to two and slight adjustments were made to question wording. Finally, for the objective familiarity item the frequency scale was replaced by on open response measurement of TV usage to reduce systematic response bias (Schwarz et al., 2008, p. 27). The adjusted items (in Dutch) for all these constructs can be found in Appendix 6.1.

For objective expertise of LCD TVs no measurement scale is described in literature. As previously discussed an objective expertise measurement should be specifically tailored to the objectives of the study (which dependent variables and in which context) (Brucks, 1985; Alba and Hutchinson, 1987). However, there are several authors who discuss an objective expertise measurement for a CE product, e.g. for a digital camera (Cordell, 1997) and for a computer (Arning & Ziefle, 2008). Inferences can be drawn from these studies for the development of an objective expertise measurement for LCD TVs. As discussed by Brucks (1985), an objective expertise measurement can consist of items on terminology, available attributes, criteria for evaluating attributes, attribute covariation and on usage situations. Although Brucks (1985) argues for an elaborate measurement of all five aspects (with open response and multiple choice items), the applied objective expertise measurements in literature consist usually of around 10-20 items in multiple choice format (Arning & ziefle, 2008; Cordell, 1997; Somasundaram, 1993) which would also fit methodological requirements in terms of questionnaire length and load on the respondents. The set of response alternatives usually contains one correct answer, three distracters and an option to indicate that the respondent does not think to know the correct answer. Subsequently, to obtain an objective expertise measurement of a respondent, the number of correct answers is counted.

To set-up an objective expertise measurement of LCD TVs the following steps were taken:

1. Development of items based on previous research, product manuals, product information on the Internet, DTV systems guides (e.g. Fischer (2004)) and input from several DTV system experts.
2. Short pilot test of initial item pool.

3. Discussion with DTV system experts and researchers on which items to include in the measurement scale.
4. Large scale pilot test of the complete measurement scale (part of full-scale pilot experiment further discussed in section 6.2.8).
5. Final adjustment of the scale.

After following this procedure, 11 items (in Dutch) were included in the final objective expertise measurement. This scale included five multiple choice, 2 check-all-that-apply and four true/false items measuring knowledge of LCD TV terminology, usage situations (including common failures in LCD TV picture quality) and technical functioning. An example of two different items is shown in Table 6.1[19]. The complete set of items and response alternatives is shown in Appendix 6.1.

Table 6.1 Example of two objective expertise items on LCD TVs (translated)

Objective expertise item	Response alternatives (correct answer in bold)
What does the abbreviation "LCD" stand for in the term "LCD television"?	Multiple choice: A. Led Coordination Display **B. Liquid Crystal Display** C. Living Color Display D. Light Compact Display E. I don't know
Red colored horizontal or vertical lines on the display of an LCD television are usually caused by defect pixels	A. Yes **B. No** C. I don't know

6.2.3 Research variables: Selection and design of the failure scenarios

As discussed in section 6.1, the goal of this study was to evaluate how consumer knowledge affects failure attribution for two different failures in the picture quality of an LCD TV: one failure caused by a fault in product's software itself and one failure caused by something external to the LCD TV. To select two relevant, realistic, distinctly different (in terms of physical cause) but also "equal" failures (in terms of objectively determined failure impact, degradation of picture quality etc.) input was used from DTV system experts in two brainstorm sessions. The following two failures were chosen:

- A failure that is most likely to be caused by (software) faults in the TV: Blocking artefacts on the TV screen.
- A failure that is most likely to be caused by a signal disturbance in the cable or a bad cable (connection): Noise on the TV screen.

[19] Please note that this example is based on a translation of the original Dutch items by the researcher and that for use of the measurement scale in other countries the items need to be properly translated and/or adjusted.

For both failure scenarios a written introduction text was shown which included a description of the basic set-up of the TV (i.e. analogue cable signal and no additional equipment such as a DVD player) and information on the conditions in which the failure scenario occurred. Furthermore, for both failures, two video-based failure scenarios were designed that had a similar introduction to the scenario (i.e. living room context and similar introduction text), similar TV content, similar duration of the failure etc. The complete failure scenario selection, design, review and pretesting process are presented in a separate study[20]. A snapshot of both failures in the failure scenario used in the experiment is shown in Figure 6.2[21].

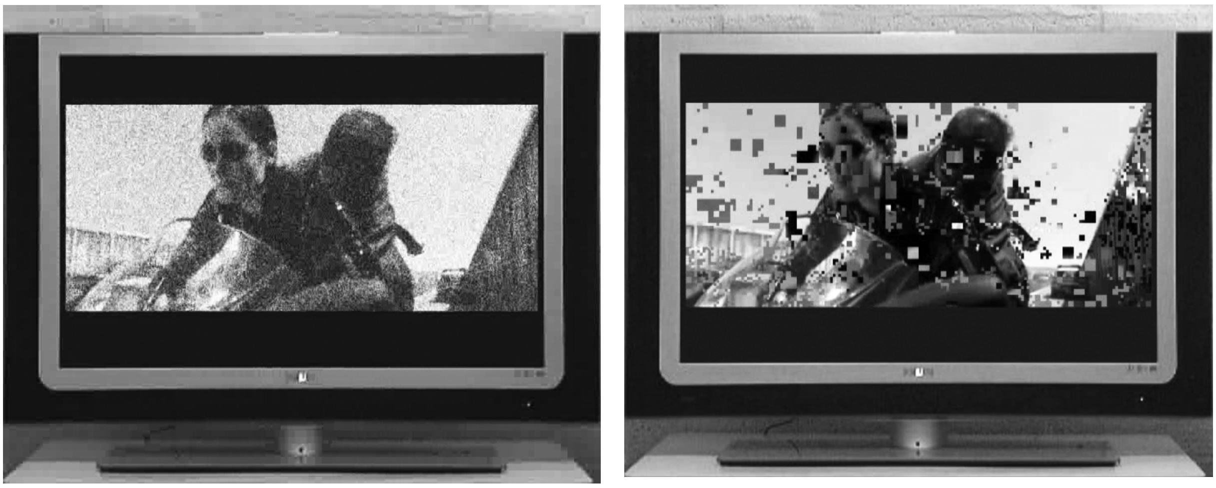


Figure 6.2 Snapshot of failure scenario with noise (left) and blocking artefacts (right)

6.2.4 Research variables: Failure attribution measurements

As shown in the conceptual framework of this study as discussed in section 6.1.4, both the three attribution dimensions and open response attribution were included as dependent variables. To measure the three attribution dimensions (i.e. locus, controllability and stability, see section 3.3) and the open response attribution defined by Oliver (1996), an adjusted version of the causal dimension scale of Russell (1982; 1987) was used. Taking the specific context of this study into account, the original items were translated into Dutch and slightly modified:

- For all items the item wording was more specified towards a failure (instead of a general outcome) and specified towards TVs and TV quality.
- For locus the item wording was changed to distinguish between a perceived cause internal to the TV versus a perceived cause outside the TV (instead of internal to the person or external to product). This change was made because for these failure

[20] This study is published in: "Keijzers, J., Scholten, L., Lu, Y. & Den Ouden, P.H. (2009). Scenario-based evaluation of perception of picture quality failures in LCD televisions. In R. Roy & E. Shebab (Eds.), *Proceedings of the 19th CIRP Design Conference.* (pp. 497–503). Cranfield: Cranfield University Press".
[21] The failure scenario videos can be obtained from the author.

scenarios the participants were asked to respond to simulated failures instead of a real experienced failure in which a participant could attribute a failure to him/herself.

- For controllability one item (i.e. outcome was intended by me or other people versus outcome was not intended by me or other people) was removed because it did not fit with the specific context of this study.

An overview of the failure attribution measurements (in Dutch) is shown in Appendix 6.1.

6.2.5 Research variables: Control variables measurements

In this section, the measurements of the control variables age, product involvement, product expectation, perceived failure impact and perceived picture quality are discussed. First of all, product involvement can be measured by several different measurement scales ranging from task or function specific involvement with one item (Lazar, Jones & Schneiderman, 2006), to product importance ratings of four items (Lancellotti, 2004), to the Personal Involvement Inventory which has 20 items (Zaichkowsky, 1985). To keep the number of items as low as possible while retaining the possibility to analyze scale validity and reliability, an adjusted and translated version of the product involvement scale developed by Lancellotti (2004, p. 242) was used. This scale consisted of three items measuring to what extent LCD TVs were important, useful and appealing to consumers on a five-point Likert scale. Similarly, for product expectations an adjusted and translated version of the product expectations scale developed by Lancellotti (2004, p. 204) was used. This scale consisted of three items measuring to what extent consumer's expected an LCD TV to function reliably, flawlessly and with a high picture quality (again measured with a five-point Likert scale).

Finally, failure impact was measured with a numerical item adjusted from De Visser (2008, p. 195) and perceived picture quality was measured on a five-point Likert scale ranging from very good to very bad. To test the validity of the design of the (failure) scenarios, measurements of failure perception (i.e. did the participants indeed observe a picture quality failure) and perceived scenario realism were included. A detailed overview of all the measurements and its items (in Dutch) discussed in this section is shown in Appendix 6.1. Next to these measurements, an open response measurement of the year of birth was included to measure the participant's age.

6.2.6 Population, sample and sampling method

Similar to the survey discussed in Chapter 4, for this web-based experiment the population of interest is a preferably heterogeneous population of Dutch consumers who are willing to use an LCD TV and who meet generally used demographic criteria such as aged 16 years old or above. To counter the disadvantages of convenience sampling used in Chapter 4, for the web-based experiment an online consumer panel and advertisements on various Internet forums were used to attract a large and heterogeneous group of respondents. This multiple-site entry technique potentially reduces self-selection bias for Internet-based questionnaires (Reips, 2002a). For access to the consumer panel members, a banner was placed on the consumer panel website for which the panel members earned €0.03 when they accessed the website linked through the banner. After 4000 referrals to experiment's website, the banner on the consumer panel website was removed. Please note that although this ensured that at least 4000 people would access the website's experiment, it was still up to the person to fill in the questionnaire or not. An overview of the response rate and distribution of respondents in terms of referring URLs is further discussed in section 6.3.1.

6.2.7 Design of Internet-based experiment

This section discusses the most important aspects of the design of the web-based experiment.

Technical aspects of the web-based experiment
The web-based experiment was designed by using Limesurvey (Limesurvey v1.71) installed on a local server on the university campus. To ensure visibility of the questionnaire a special website address referring to TV quality was created (www.televisiekwaliteit.id.tue.nl). The video-based failure scenarios were embedded in the questionnaire by using YouTube (YouTube). Since research has shown that Internet-based experiments lessen the degree of control of the experimental setting, additional data on the responses was logged (Reips, 2002a) First of all, to control for multiple entries by the same person the IP address and referring URL of the respondent was logged. Secondly, the time it took respondents to complete the experiment was logged to control for erroneous answers. Respondents taking either too little or too much time to complete the questionnaire should be removed before further analysis due to potential threats of experimental validity.

Experimental procedure
This section briefly discusses the experimental procedure and ordering of the questionnaire items. Besides following the general guidelines on questionnaire, question and item construction (Dillman, 2000; Fowler & Cosenza, 2008; Manfreda & Vehovar, 2008; Schwarz et al., 2008) (see also section 4.2.3), special attention was paid to reduce the drawbacks of Internet-based experiments (Reips, 2002a). An overview of the complete experimental procedure followed by each participant is shown in Figure 6.3. All the items of the questionnaire and the introduction text of the failure scenario can be found in Appendix 6.1.

From this figure can be seen that all participants were shown a video of the LCD TV without an implemented failure in the beginning of the questionnaire to create a similar frame of reference for all participants (i.e. evaluating picture quality of an LCD TV via an Internet-based experiment with video-based scenarios). Furthermore, each participant was asked to rate the picture quality of the LCD TV shown in this scenario and to indicate whether they had seen a failure in this scenario in order to test: 1) significant difference with the failure scenario; and 2) for potential confounding (computer or Internet related) variables such as Internet browser or network problems (Reips, 2002a). Similarly, also for the failure scenario a control item on failure perception was included to ensure that every participant did observe the product failure as intended. To ensure random allocation to either the noise or blocking artefacts scenario, the participants' year of birth (even or uneven) was used as a condition for referral.

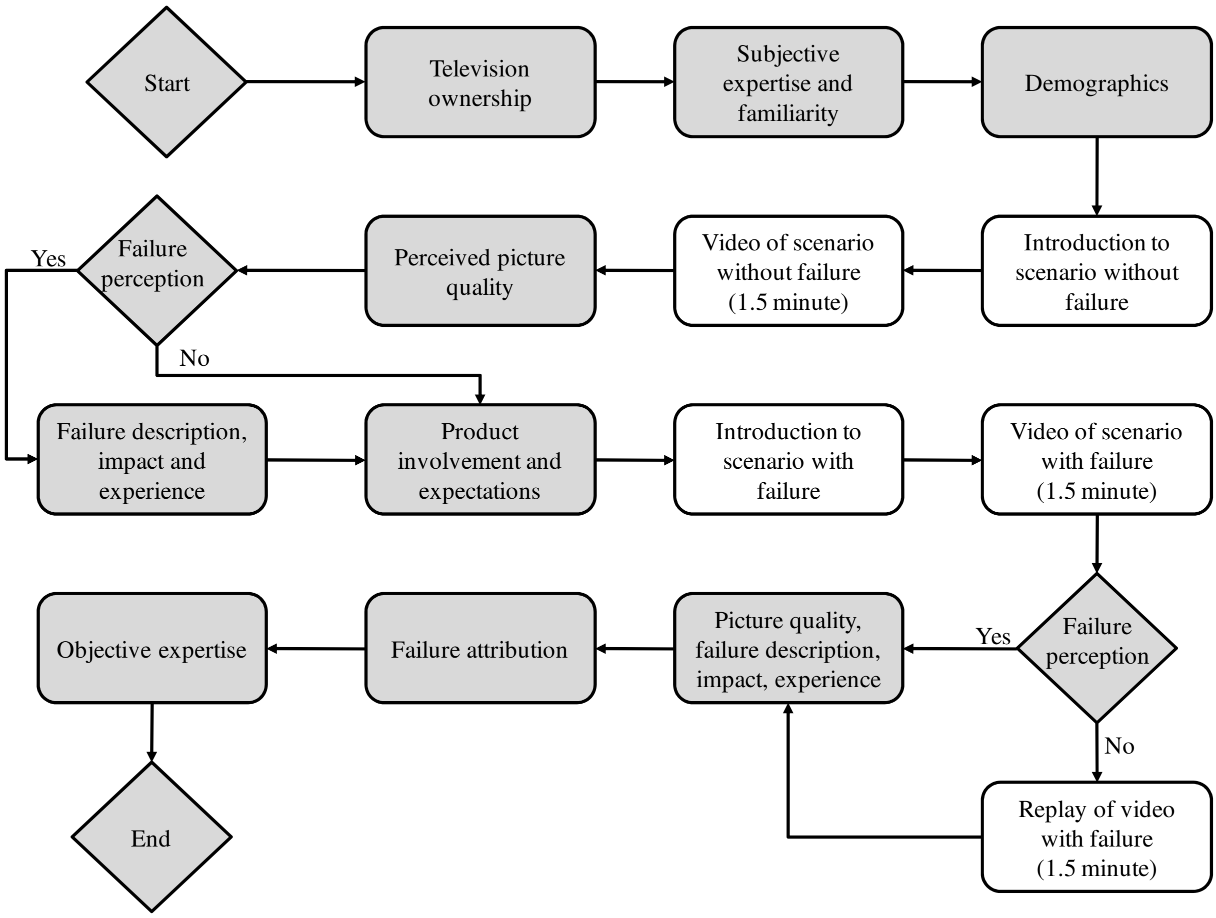


Figure 6.3 Overview of the procedure of the Internet-based experiment

Similar to the survey discussed in Chapter 4, several potential order effects were taken into account when ordering these subjects in the web-based experiment (Dillman, 2000, p. 89; Reips, 2002a). First, the measurement of subjective expertise was put *before* the failure scenarios and objective expertise measurements because of a potential carryover effect. For the same reason and to improve response rate objective expertise items were put at the end of the questionnaire due to potential carry-over effects to the attribution measurement. Secondly, product involvement and product expectations were measured *after* a video of the LCD TV without a failure (to create a similar frame of reference for each participant) and *before* the failure scenario (to prevent potential bias due to watching a product failure).

Methods to improve response rate

To improve the response rate (besides the advantage of using a (paid) consumer panel) several methods were used of which the most important are (Dillman, 2000; Lynn, 2008; Reips, 2002a):

- Ensuring confidentiality of information provided.
- Use of the university and research project name and logo to emphasize the importance and professionalism of the research project.
- Price draw of three times €100,- in cash among the respondents who completed the web-based experiment and met the inclusion criteria (i.e. age, no double responses etc.).
- Email address at every page of the web-based experiment to reach the researcher in case of problems.

6.2.8 Pilot experiment

Before proceeding with the large scale experiment, the questionnaire design and appearance on different computers with different Internet browsers and Internet connections (Reips, 2002a) was tested in a small-scale pilot test (n = 15) among colleagues and family. Based on the results small problems with questionnaire readability on difference Internet browsers as well as the visibility of the embedded videos were solved.

Subsequently, the complete experiment was pre-tested in a pilot experiment with 40 participants (students of the faculty Industrial Design at Eindhoven University of Technology). The set-up and results (in terms of selection of scenario content and tests of scenario realism) of this pilot experiment are discussed in a separate study presented in a separate paper (Keijzers, Scholten, Lu & Den Ouden, 2009). Based upon the results of this experiment two objective expertise items were replaced (due to low validity), the introduction text of the failure scenarios was improved and small adjustments were made to item wording.

6.3 Results

In this section, the results of the experiment are discussed. First, in section 6.3.1 an overview is given of the characteristics of the respondents. Subsequently, section 6.3.2 discusses the validation of the different constructs and measurements scales used throughout the experiment. Based on these results section 6.3.3 discusses the set-up and assumptions check of the statistical analyses and coding of the open response attribution measurement to test the hypotheses. Next, in section 6.3.4, the overall effect of consumer knowledge and failure origin on failure attribution is discussed. Finally, to gain deeper insight into failure attribution differences, in sections 6.3.5 and 6.3.6 the scenario-specific results and the effect of consumer knowledge is discussed for respectively the noise and blocking artefacts scenario.

6.3.1 Respondent characteristics

In this section an overview is given of the response rate and respondent characteristics. The first entry was recorded on June 6th 2008 and the final entry on June 30th 2008. In total 657 responses were recorded out of which 408 were fully completed questionnaires. This large difference between recorded responses and fully completed questionnaires was due to the fact that any press of a button in the questionnaire would result in a record of the response. Most incomplete responses were from participants who started the questionnaire but did not answer any question. The results of the logging of time showed it took on average slightly less than 13 minutes to complete the questionnaire. Based on these results and results of the pilot experiments, responses which took either less than five minutes or more than 45 minutes to complete the experiment were excluded from further analysis. After further excluding responses which did not meet the other inclusion criteria (e.g. open response answers not related to the content of the questionnaire, description of the failure scenario which did not match the content, multiple entries from the same IP address), 354 remained for further analysis.

An overview of the distribution among referring URLs of these respondents is shown in Table 6.2. From this table can be seen that more than 80% of the respondents were recruited via the consumer panel. Next, an overview of the respondent characteristics in terms of age, educational level and gender is shown in Table 6.3. From this table can be seen that participants ranged from 16 to 65 years old and were mostly medium to highly educated (89.5%). Furthermore, a slight bias towards female participants was observed due to the use of the consumer panel.

Table 6.2 Overview of distribution of respondents among referring URLs

Referring URL		n	%
Moneymiljonair	www.moneymiljonair.nl	295	83.3
Tweakers	www.tweakers.net	6	1.7
Elektronicaforum	www.elektronicaforum.nl	5	1.4
Thesistools	www.thesistools.nl	5	1.4
Scholieren.com	www.scholieren.com	6	1.7
DVDforum	www.dvdforum.nl	8	2.3
Computertotaal	www.computertotaal.nl	8	2.3
HTforum	www.htforum.nl	12	3.4
50plusplein	www.50plusplein.nl	1	0.3
Unknown	-	8	2.3
Total	**-**	**354**	**100.0**

Out of the 354 respondents, 344 owned a TV out of which in turn 132 owned a plasma and/or LCD TV. Furthermore, 50.6 % of the respondents used a digital cable signal.

Table 6.3 Overview of respondent characteristics in terms of age, educational level and gender

Age	n	%	Educational level	n	%	Gender	n	%
< 21 years	33	9.3	Low	37	10.5	Male	148	41.8
21 – 30 years	94	26.6	Medium	210	59.3	Female	206	58.2
31 – 40 years	89	25.2	High	106	29.9			
41 – 50 years	78	22.0						
51 – 60 years	51	14.4						
61 – 70 years	9	2.5						
71 – 80 years	0	0.0						
80 > years	0	0.0						
Missing	0	0.0		1	0.3		0	0.0
Total	**354**	**100.0**		**354**	**100.0**		**354**	**100.0**

6.3.2 Validation of the measurements

In this section the validation of the measurements for the independent, dependent and control variables is discussed. To analyze the reliability and validity of the objective expertise measurement, first the scores on the individual items were calculated. Each item was given the same weight in the total score. Only the correct answer was rewarded with one point. For the "check-all-that-apply" questions first a score was obtained by counting the number of correctly selected answers and deducting the number of incorrect answers. When this score was equal to or higher than the total number of possible correct answers deducted by two (i.e. approximately half of the answers correct), the participants received a point for that item. For example, for objective expertise item number six this score had to be equal to or higher than two (i.e. this items had four correct answers) to be awarded one point for this item.

To assess the reliability and validity of the objective expertise measurement, the point-biserial correlation, p-values and Cronbach's alpha were used (DIIA, 2003; Varma, n.d.). In this context the point-biserial correlation can be used to assess item quality because it is an indication of the discriminatory power of an item (Varma, n.d.). Consequently, items with a point-biserial value below 0.15 (Varma, n.d.) need to be removed from further analysis. Furthermore, the p-value can be used as an indicator of item difficulty. Items with a p-value above 0.90 (too easy) or a p-value below 0.20 (too difficult) should be considered for removal from further analysis (DIIA, 2003). From the results of this analysis shown in Appendix 6.2 can be seen that all the items are valid for further use in the analysis. Furthermore, the results of the analysis indicate the complete scale has a Cronbach's alpha of 0.808 which is very good for this type of scale (DIIA, 2003). Overall, the results showed a mean objective expertise score of 5.07 (S.D. = 3.00) and, although not normally distributed, a reasonable equal spread among the score as shown in Figure 6.4.

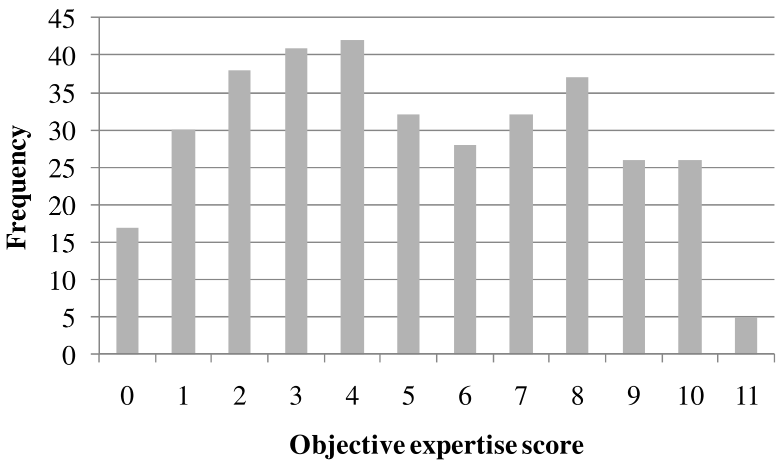


Figure 6.4 Distribution of objective expertise score

Since the different consumer knowledge constructs were measured on different scales, they could not be combined into a single factor analysis as was done with the measurements in Chapter 4. The results of the separate factor analyses for subjective expertise and subjective familiarity show that the items scored on a single factor and measurements of Cronbach's alpha showed that both scales were reliable (0.877 for subjective expertise and 0.820 for subjective familiarity).

For the factor analysis of the causal dimension scale mixed results were found. The results of factor analyses showed that the controllability items did not score on a single factor and had to be removed from further analysis. Furthermore, although the stability construct did emerge from the factor analysis, separate measurements of Cronbach's alpha (0.881 for locus and 0.320 for stability) indicated that the scale was not reliable and therefore not suitable for further analysis. The locus scale which was of main interest for this study proved to be both valid and reliable. When reflecting on the wording of the stability and controllability items, a possible explanation for the weak validity and reliability could be that the items were too difficult to answer due to the use of failure scenarios instead of a real-life failure experience.

Next, results of the factor analyses for the product involvement and expectations measurement scales show that both scales are valid. Separate measurements of Cronbach's alpha proved that both scales are reliable (0.826 for product involvement and 0.806 for product expectations) and therefore suitable for further analysis. A summary of the descriptive statistics of the constructs discussed in this section is shown in Table 6.4. Please note that for both the product involvement and product expectations a reverse scored measurement is used so that a positive score refers to a higher level of involvement or higher (more positive) expectations.

Table 6.4 Descriptive statistics for the questionnaire constructs

Construct	Mean	S.D.	Scale range	Number of items
Objective expertise	5.07	3.00	0 – 11	11
Subjective expertise	2.75	1.24	1 – 5	2
Objective familiarity	23.96	17.57	0 – 105	1
Subjective familiarity	3.64	1.10	1 – 5	2
Product involvement (reverse scored)	3.41	0.95	1 – 5	3
Product expectations (reverse scored)	4.35	0.72	1 – 5	3

From this table can be seen that for objective expertise the mean score is slightly below the mean of the scale while for subjective expertise and subjective familiarity the mean score is above the mean of the scale. Furthermore, for product involvement and particularly for product expectations a high mean score was observed. In other words, participants were on average interested in LCD TVs and had high expectations concerning the product's quality and reliability.

By using the mean score of the expertise and familiarity constructs, both convergent and discriminant validity can be discussed by investigating the correlations between these constructs (see also Chapter 4). An overview of the correlations between the consumer knowledge constructs and the control variables is shown in Table 6.5. When comparing these results with the results of the survey discussed in Chapter 4, similar effects could be observed except from subjective familiarity. Due to shortening of the scale and only focusing on items reflecting usage, only a significant correlation with objective familiarity was observed. Moreover, since objective expertise neither (positively) correlates with both familiarity constructs, this confirms the conclusions drawn in Chapter 4 and 5 that usage experience of TVs is generally passive and does not automatically result in an increase of (either perceived or objective) expertise. As can be seen in the literature review by Carlson et al. (2009)[22] on the relation between objective and subjective expertise, the relatively high correlation between objective and subjective expertise of TVs found in this study is of the same magnitude for other CE or technological products.

Finally, from the table can be seen that for the product involvement and product expectations measurements, significant correlations with all consumer knowledge constructs were observed. In other words, consumer knowledge on LCD TVs is positively related to interest in LCD TVs and higher expectations of LCD TV quality and reliability.

[22] Please note that this information was published after conducting this experiment and was therefore not available for previous studies discussed in this dissertation.

Table 6.5 Correlations (Spearman's rho) of questionnaire constructs, N = 354

	Subjective expertise	Objective familiarity	Subjective familiarity	Age	Product involvement (reversed)	Product expectations (reversed)
Objective expertise	0.591**	-0.105*	-0.038	0.041	0.227**	0.173**
Subjective expertise		-0.095	0.013	-0.116*	0.315**	0.119*
Objective familiarity			0.550**	0.210**	0.222**	0.131*
Subjective familiarity				0.021	0.329**	0.278**
Age					-0.074	0.119*
Product involvement						0.475**

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).

6.3.3 Set-up of the analyses

To investigate the main effect of consumer knowledge on failure attribution, multiple separate MANOVAs (for each consumer knowledge construct separately) with a 2 (consumer knowledge) x 2 (failure cause) factorial design were used in which the attribution locus scale, perceived failure impact and perceived picture quality were included as dependent variables. Similar to the method used in Chapter 4, participants were categorized into low versus high consumer knowledge based on a split on the mean value (see Table 6.4) of the different consumer knowledge constructs. An overview of the characteristics of both groups based on a differentiation on objective expertise is shown in Table 6.6. Furthermore, results of the assumptions check for the MANOVA analysis for objective expertise showed that the data sufficiently met the criteria for performing this analysis.

For the analysis of the open response attribution measurement all the responses were categorized into groups of similar attributions in a discussion session with three DTV system experts. Subsequently, the effect of consumer knowledge on failure attribution in terms of differences in attributed causes was investigated by using nonparametric separate pair-wise Mann-Whitney U tests. For both the MANOVA and the nonparametric tests the level of significance is set at $p = 0.05$. Results within the less restrictive level of $p = 0.1$ are indicated as marginally significant. Finally, for the MANOVA analyses the significance of the omnibus F-tests were taken from the Pillai values.

Table 6.6 Overview of participant characteristics based on a differentiation on objective expertise

	High objective expertise			Low objective expertise		
	Mean	**S.D.**	**Range**	**Mean**	**S.D.**	**Range**
Objective expertise	8.04	1.45	6-11	2.76	1.53	0-5
Subjective expertise	3.44	1.20	1-5	2.19	0.95	1-5
Objective familiarity	22.36	16.53	1-105	25.20	18.27	0 - 105
Subjective familiarity	3.54	1.14	1-5	3.70	1.06	1-5
Age (years)	36.67	12.01	16-65	36.88	12.48	18-65
Product involvement	3.58	1.01	1-5	3.28	0.87	1-5
Product expectations	4.39	0.64	1-3.67	4.31	0.81	1-5

6.3.4 Evaluation of the overall effect of consumer knowledge and failure cause

In this section, the overall effect of consumer knowledge and failure cause on the dependent variables is discussed as well as the results for the evaluation of control variables (effect of age, product involvement and product expectations) and possible confounding factors (effect of failure experience).

Evaluation of control measurements of failure perception and perceived scenario realism

Before proceeding with the analysis of the main effects, the design of the experiment was evaluated. First, a Wilcoxon signed-rank test was used to test whether the perceived picture quality scores of the introduction scenario (without failure) and the failure scenario significantly differed. The results confirmed that the picture quality of the LCD TV in the failure scenarios was perceived as significantly worse than the picture quality of the LCD TV in the introduction scenario ($p < 0.001$). Secondly, a separate pair-wise Mann-Whitney U test confirmed that there was no significant difference in perceived scenario realism between the two failure scenarios ($p < 0.8$). For both scenarios the mean score of perceived scenario realism (2.60 for the noise scenario and 2.58 for the blocking artefacts scenario) showed that both scenarios were perceived, on average, as moderately realistic (see Appendix 6.1 for the measurement scale used).

Furthermore, separate pair-wise Mann-Whitney U tests confirmed that there were no significant differences between both failure scenarios in terms of perception of the presence of a failure in the failure scenario ($p < 0.2$) and in terms of previous experience with the failure ($p < 0.2$). For both failure scenarios approximately 80% of the respondents immediately perceived the occurrence of a failure in the failure scenario. For the noise scenario 43% of the respondents indicated having experienced this type of failure in the past in comparison to 51% for the blocking artefacts scenario. Summarizing can be concluded that the design of the failure scenarios was successful.

MANOVA results for the effect of consumer knowledge and failure cause

To evaluate the effect of objective expertise and failure cause on failure attribution, several MANOVA models were used to evaluate the effect of the moderating and control variables as shown in the conceptual research framework in Figure 6.1. Since the effect of product involvement and product expectations as blocking factors was not significant and did not improve the significance and power for the main independent variables, these were stepwise removed from the model (Hair et al., 2006, p. 419). Consequently, the final MANOVA model incorporated, besides the main independent variables, age as a covariate and failure experience as a blocking factor. The results of the multivariate tests and tests of between-subjects effects for this model are shown in Appendix 6.3. From this model can be seen that the overall effect of objective expertise ($F (3, 343) = 7.512, p < 0.001$), failure experience ($F (3, 343) = 3.645, p < 0.05$) and age ($F (3, 343) = 2.667, p < 0.05$) on the dependent variables was significant. However, the results also showed that the overall effect of failure cause was not significant ($F (3, 343) = 1.535, p < 0.3$). In other words, contrary to what was expected, the cause of the failure did not have a significant effect on failure attribution, perceived picture quality and perceived failure impact. Finally, no significant interactions between the dependent variables and control variables were found.

Further tests of the between-subjects effects showed that objective expertise had a significant effect on attribution locus ($F (1, 354) = 19.007, p < 0.001$) and perceived picture quality ($F (1, 354) = 5.162, p < 0.05$) but not on failure impact ($F (1, 354) = 0.158, p < 0.7$). Based on this analysis can be concluded that for objective expertise the effect on perceived picture quality is significant and therefore hypothesis H_5 needs to be accepted. For failure impact the MANOVA results for objective expertise showed no significant difference and therefore H_6 needs to be rejected.

From the interaction between objective expertise and failure cause for attribution locus shown in Figure 6.5 can be seen that both failures were on average perceived to be caused more by TV external factors than TV internal factors (mean of attribution scale is three). For both scenarios higher levels of objective expertise also resulted in a more extreme external attribution (i.e. attributions towards a cause outside the TV) compared with lower levels of objective expertise. Specific results for each failure scenario and the results of the hypothesis tests for attribution locus are further discussed in section 6.3.5 and 6.3.6.

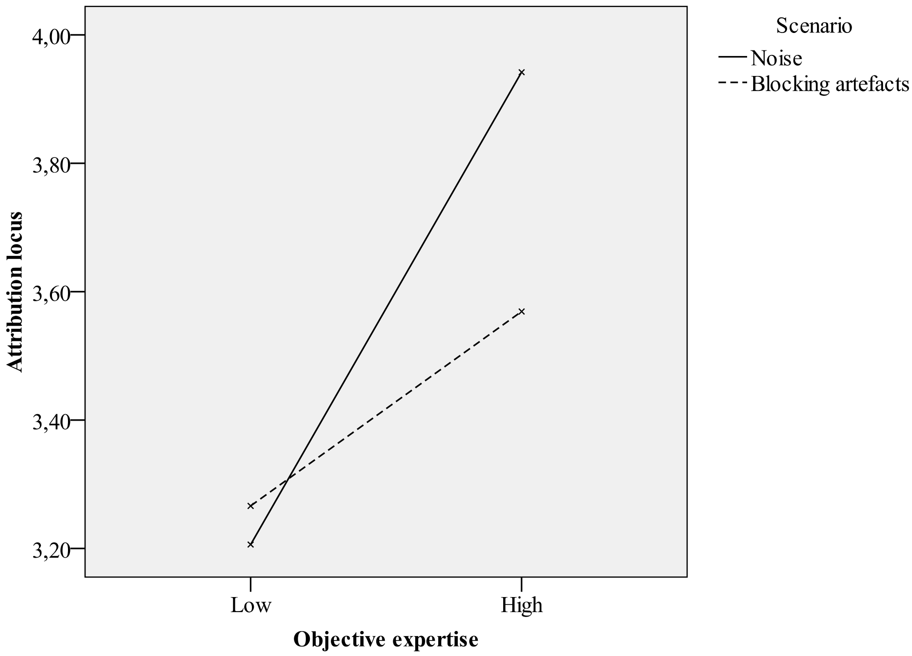


Figure 6.5 Interaction plot of objective expertise and failure cause for attribution locus (higher scores refer to more external attributions)

Furthermore, from the interaction plot between objective expertise and failure cause for perceived picture quality shown in Figure 6.6 can be seen that for both scenarios higher levels of objective expertise also resulted in a more negative judgment of perceived picture quality compared with lower levels of objective expertise.

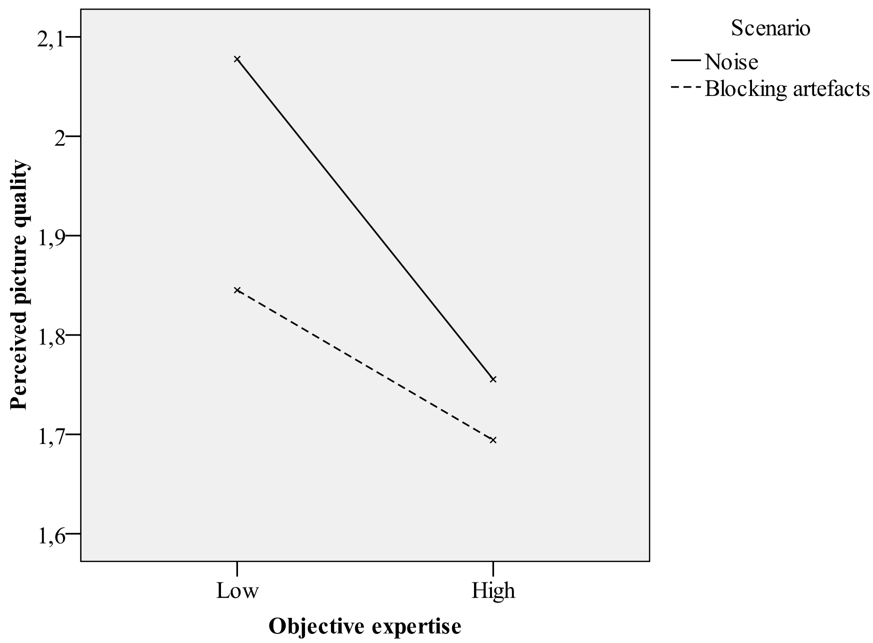


Figure 6.6 Interaction plot of objective expertise and failure cause for perceived picture quality

Separate tests of the between-subjects effects for failure experience showed the effect of this variable was significant for attribution locus ($F(1, 354) = 4.323$, $p < 0.05$) and perceived picture quality ($F(1, 354) = 6.551$, $p < 0.05$) but not on failure impact ($F(1, 354) = 0.045$, $p < 0.9$). Interaction plots of failure experience and failure cause for attribution locus and perceived picture quality are shown in Figure 6.7 and Figure 6.8 respectively.

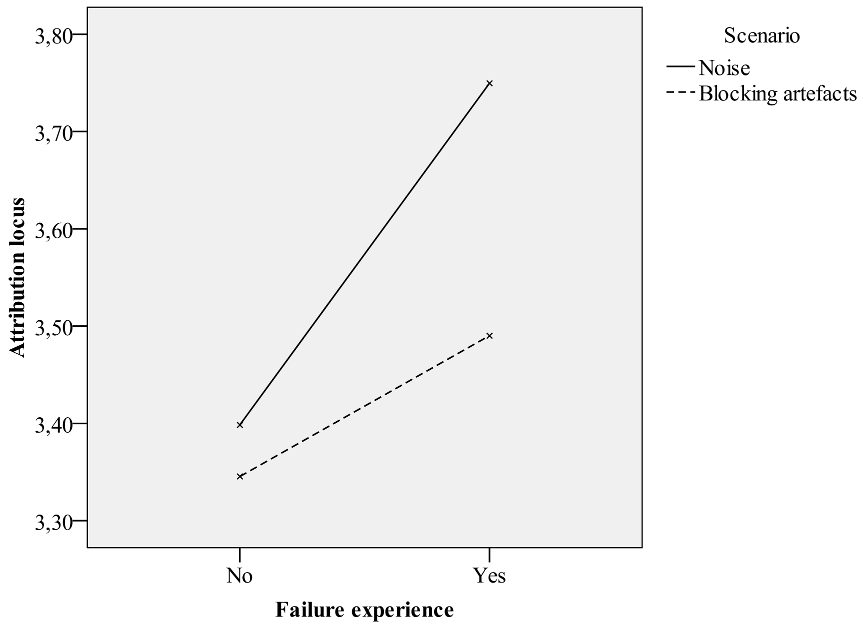


Figure 6.7 Interaction plot of failure experience and failure cause for attribution locus (higher scores refer to more external attributions)

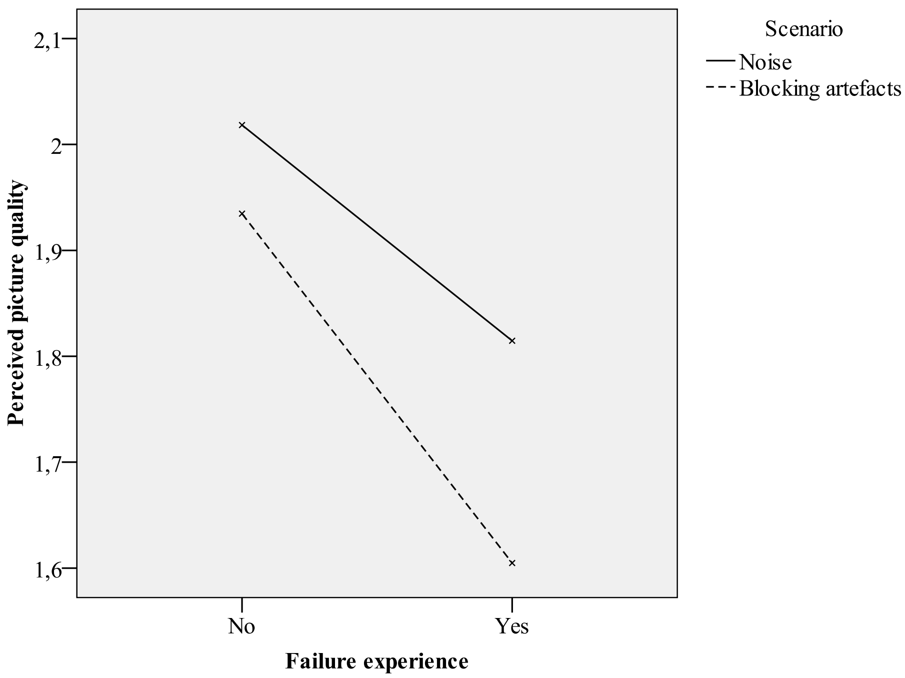


Figure 6.8 Interaction plot of failure experience and failure cause for perceived picture quality

From both figures can be concluded that experience with a failure had a similar, but significantly smaller, effect as objective expertise. For both scenarios experience with a failure resulted in a more extreme attribution towards external causes and a more negative evaluation of perceived picture quality.

The results of similar separate MANOVAs with subjective expertise ($F(3, 343) = 1.181$, $p < 0.4$), objective familiarity ($F(3, 343) = 1.138$, $p < 0.4$) and subjective familiarity ($F(3, 343) = 0.735$, $p < 0.6$) as main independent variable indicated that the overall effect of these consumer knowledge constructs on the dependent variables was not significant[23]. Hypothesis H_4, which stated that objective expertise stronger relates to differences in failure attribution than subjective expertise and familiarity, can therefore be accepted.

Evaluation of the overall effect of consumer knowledge and failure cause on the open response attribution measurement

Finally, the overall results of the open response attribution measurement are discussed. During a discussion session with DTV system experts the individual attribution responses were subdivided into one or more differently perceived causes. Based on agreement between the experts similar causes were subsequently grouped and labeled. The overall results of this coding process compared for both scenarios are shown in Figure 6.9 and Figure 6.10. Figure 6.9 shows a differentiation between attribution to one or multiple causes inside the TV (internal), to one or more causes outside the TV (external) and to mixed causes (both internal and external). Figure 6.10 gives an overview of the relative frequencies with which the differently labeled causes were mentioned. For example, “External-signal” refers to an attribution of the failure to the quality of the transmission of the TV signal, “External-TV settings” refers to an attribution to wrong settings selected by the user (referred to as an external attribution because the user and not the product is blamed) and an attribution to “TV plus signal” refers to a perceived mismatch between the quality of the TV and the quality and/or type of TV signal. For TV internally attributed causes no further differentiation was made since most of the responses did not further specify the answer beyond the TV in general.

Based on these results several conclusions can be made. First of all, results of a separate pair-wise Mann-Whitney U test confirmed the findings discussed above: there is no significant difference for perceived locus of the failure ($p < 0.3$). In other words, there is a difference between the consumer’s and designer’s attribution of the failures. Only 35% of the respondents attributed the blocking artefacts to (something inside) the TV while designers confirmed that this failure, as shown in the scenario, is caused by a fault in the software of the TV.

---

[23] For these analyses the effect of both failure experience as blocking factors and age as covariate were significant and similar to the model with objective expertise as independent variable.

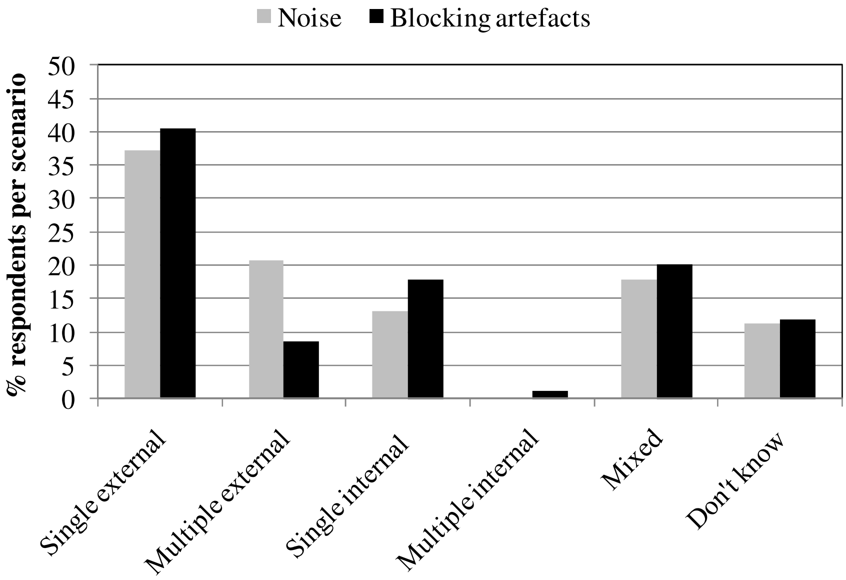


Figure 6.9 Overview of attribution locus per scenario (shown as percentage of the total number of respondents per scenario)

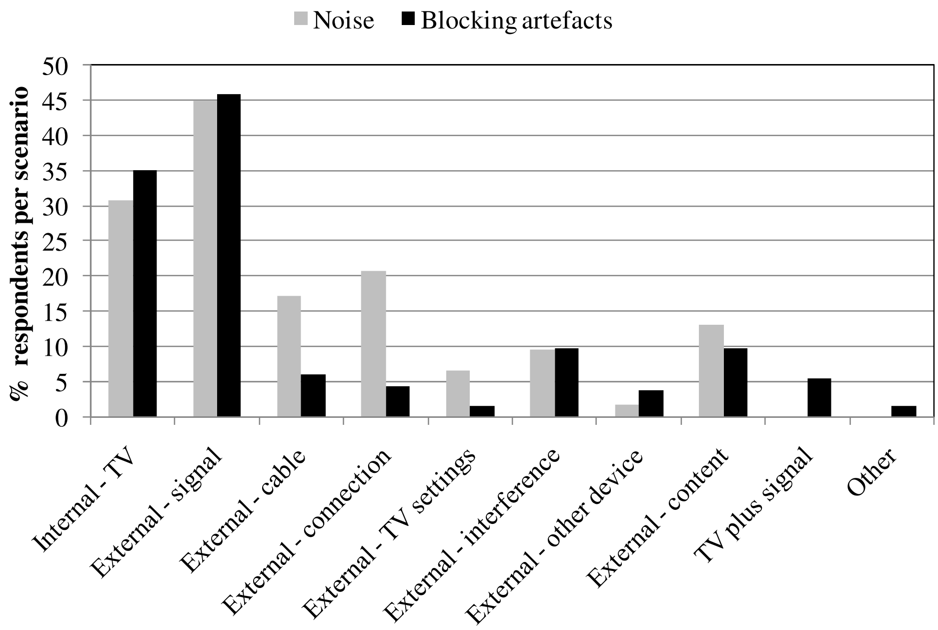


Figure 6.10 Overview of attributed causes per scenario (shown as percentage of the total number of respondents per scenario)

Although the noise scenario is attributed more in agreement with the product designers, still 30% of the respondents perceived that an internal fault in the TV itself can be the cause. Moreover, for both failure scenarios a large "spectrum" of attributed causes is mentioned. Although there is a striking number of similarities in type of attributed causes, separate pair-wise Mann-Whitney U tests showed that for the noise scenario significantly more respondents attributed the failure to the cable ($p < 0.01$), the connection ($p < 0.001$), the TV settings (p < 0.05) and a combination of the TV and type of TV signal (p < 0.01), than for the blocking artefacts scenario.

Based on the number of separately identified attributed causes deducted from the open response measurement for each participant, an overall comparison was made between the high and low levels of objective expertise. For this analysis a differentiation was made between responses referring to multiple causes (either multiple internal, multiple external or mixed), responses referring to single causes and responses which did not refer to any cause (i.e. "I do not know"). Results of separate pair-wise Mann-Whitney U tests showed that the participants in the high objective expertise group overall mentioned significantly more different causes than participants from the low objective expertise group (p < 0.001). Consequently, hypothesis H_3 needs to be accepted. The specific effect of objective expertise on the open response attribution answers is further discussed for each scenario specifically in the following sections.

6.3.5 Analysis of the results for the noise scenario

In this section, detailed results for the noise scenario are discussed. As only the MANOVA model with objective expertise showed a significant effect, for both scenarios the analysis and discussion of the effect of consumer knowledge on the dependent variables is limited to objective expertise. An overview of the descriptive statistics of the dependent variables for the noise scenario when differentiating on objective expertise is shown in Table 6.7.

Table 6.7 Descriptive statistics of dependent variables for the noise scenario

	High objective expertise			**Low objective expertise**		
	Mean	**S.D.**	**Range**	**Mean**	**S.D.**	**Range**
Failure impact	4.12	1.07	0-5	3.90	1.18	0-5
Picture quality	1.74	0.91	1-4	2.10	0.95	1-4
Scenario realism	2.63	1.21	1-5	2.58	1.00	1-5
Attribution locus (mean)	3.91	1.23	1-5	3.20	1.05	1-5

Based on these results for this scenario, a comparison was made between these measurements for high and low levels of objective expertise. Separate pair-wise Mann-Whitney U tests showed significant differences between high and low levels of objective expertise for attribution locus (p < 0.001), perceived picture quality (p < 0.01), failure impact (p < 0.05) and number of attributed causes (p < 0.01). Since the results showed that participants from the high objective expertise group attributed the noise scenario significantly stronger to external causes than participants from the low objective expertise group, hypothesis H_2 can be accepted.

Differences between the low and high levels of objective expertise for the open response attribution measurement are shown in Figure 6.11 and Figure 6.12.

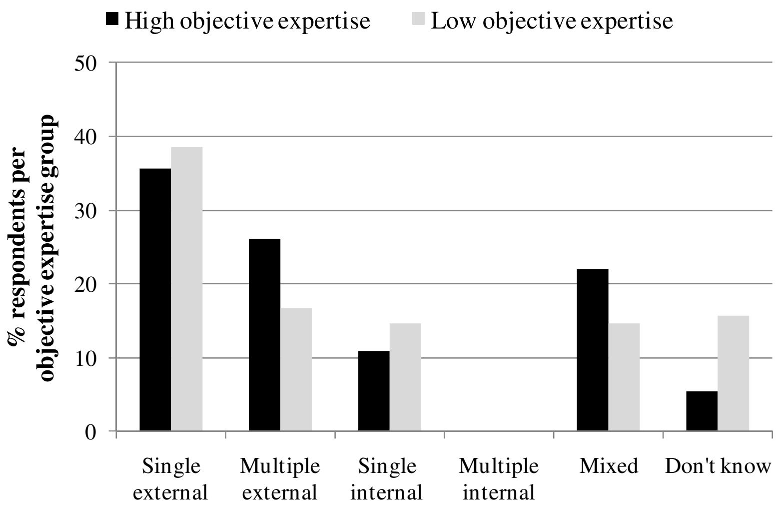


Figure 6.11 Overview of attribution locus per objective expertise group for the noise scenario (shown as percentage of the total number of respondents in the objective expertise group)

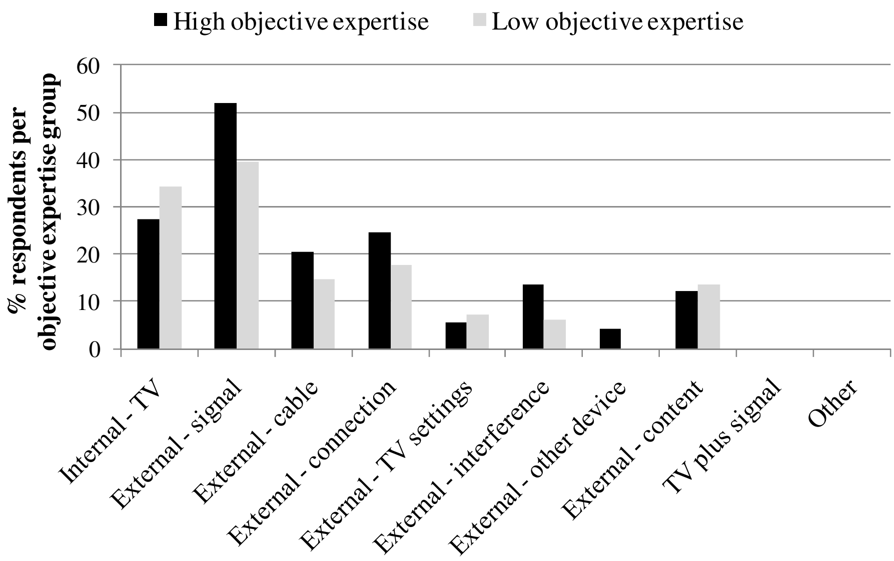


Figure 6.12 Overview of attributed causes per objective expertise group for the noise scenario (shown as percentage of the total number of respondents per objective expertise group)

From figure 6.11 can be seen that participants with higher objective expertise attribute more to multiple external and mixed causes than participants with lower levels of objective expertise who attribute more to single causes. To further illustrate this difference, below an example of two exemplary translated attribution responses is given for high and low objective expertise.

Participant in the low objective expertise group: "I think that the technical quality of a TV determines the amount of noise. In that case the manufacturer of the TV is responsible for the noise".

Participant in the high objective expertise group: "(The failure) can have different causes. It can be a disturbance at the cable signal provider, there can be atmospheric disturbances, the antenna cable is not plugged in correctly but it can also be a disturbance in the TV itself in the signal capturing part or in the image processor".

Although some of the attributed causes mentioned by this participant from the high objective expertise group are considered to be highly unlikely by the DTV system experts (i.e. disturbance in the TV), the difference between these attribution responses does show that higher objective expertise does results in more complex reasoning regarding failure causes.

6.3.6 Analysis of results for the blocking artefacts scenario

In this section, detailed results for the blocking artefacts scenario are discussed. An overview of the descriptive statistics of the dependent variables for the blocking artefacts scenario when differentiating on objective expertise is shown in Table 6.8.

Table 6.8 Descriptive statistics of dependent variables for the blocking artefacts scenario

	High objective expertise			Low objective expertise		
	Mean	**S.D.**	**Range**	**Mean**	**S.D.**	**Range**
Failure impact	4.10	1.31	0-5	4.21	1.28	0-5
Picture quality	1.69	0.98	1-4	1.86	1.05	1-5
Scenario realism	2.68	1.29	1-5	2.51	1.11	1-5
Attribution locus (mean)	3.56	1.18	1-5	3.26	1.01	1-5

Similar to the discussion of the results of noise scenario, a comparison was made between these measurements for high and low levels of objective expertise. Separate pair-wise Mann-Whitney U tests showed significant differences between high and low levels of objective expertise for attribution locus ($p < 0.05$) and number of attributed causes ($p < 0.01$), but not for perceived picture quality ($p < 0.2$) and failure impact ($p < 0.8$). Since the results show that participants from the high objective expertise group attributed the blocking artefacts scenario significantly stronger to external causes than participants from the low objective expertise group, hypothesis H_1 needs to be rejected. In other words, for this scenario a higher level of objective expertise does not result in a more correct attribution of the failure compared to a lower level of objective expertise.

Differences between the low and high levels of objective expertise for the open response attribution measurement are shown in Figure 6.13 and Figure 6.14.

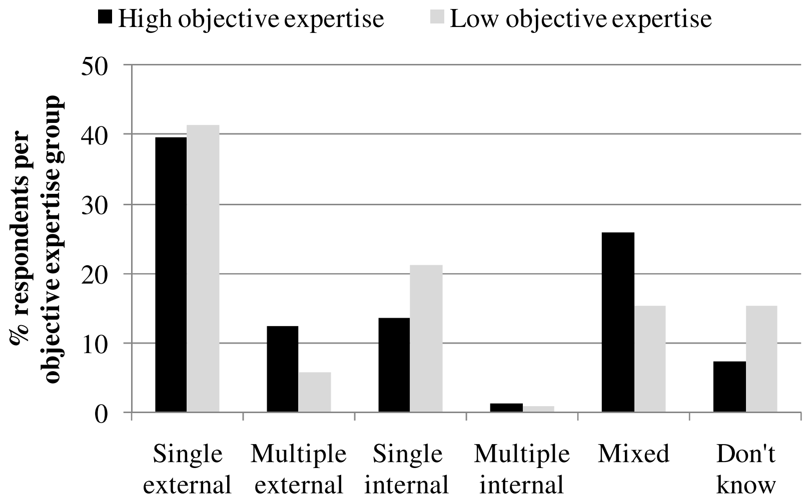


Figure 6.13 Overview of attribution locus per objective expertise group for the blocking artefacts scenario (shown as percentage of the total number of respondents objective expertise group)

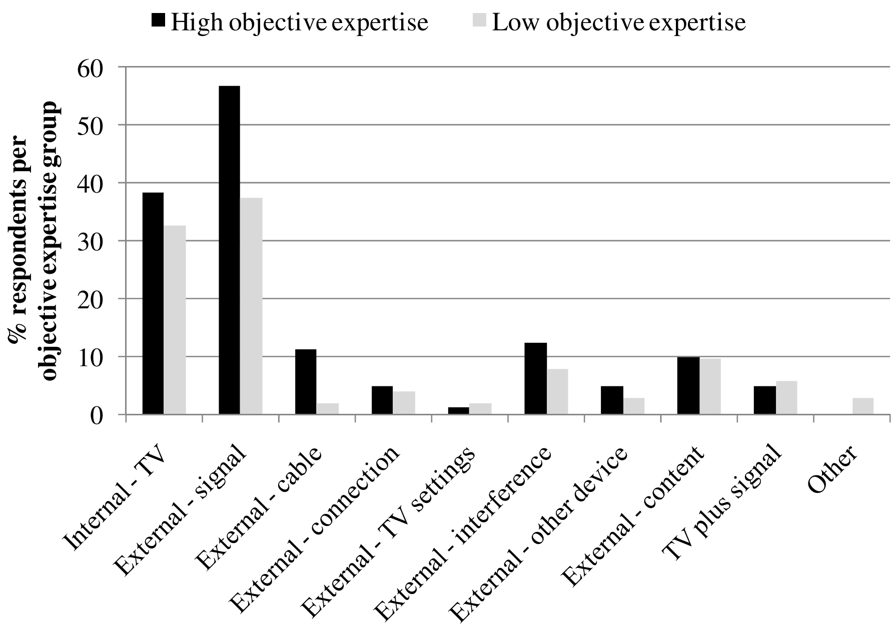


Figure 6.14 Overview of attributed causes per objective expertise group for the blocking artefacts scenario (shown as percentage of the total number of respondents per objective expertise group)

Similar to the results of the noise scenario, from figure 6.13 can be seen that participants with higher objective expertise attribute failures more to multiple external and mixed causes than participants with lower levels of objective expertise who attribute more to single causes.

Furthermore, from the results shown in figure 6.14 can be seen that almost 60% of the participants in the high objective expertise group mention the cable signal as a perceived cause of the failure. Although the designers argued that the blocking artefacts caused by software fault as shown on the designed failure scenario were distinctly different from blocking artefacts due to a bad (digital) cable signal (e.g. in frequency, appearance, severity and duration), the participants from the high objective expertise group attributed the failure to causes that are probably more familiar. For example, one participant from the high objective expertise group wrote (translated): "The cause can be determined easily: it is the source to which the TV is connected (either a DVD player, set-top box or other device). TVs do not have these types of fault and therefore cannot be blamed for this failure. (When blaming the TV) the failure should have something to do with how the image is shown: too light, too dark, stripes, ghosting, a delay, bad de-interlacing etc. Blocking artefacts such as these are mostly caused by a problem with the source. This could be damage to a DVD or a too low bitrate at the cable signal provider". In other words, for this scenario a higher level objective expertise does not result in attributions more in accordance with the DTV system experts.

6.4 Conclusion and discussion

This section concludes this chapter and discusses the results of the study and its implications. First, section 6.4.1 gives an overview of the results and discusses the hypotheses. Subsequently, section 6.4.2 discusses how the results of this chapter can help product designers to better understand the consumer's perception of and reaction to product failures. Finally, in section 6.4.3 limitations of the study discussed in this chapter and subsequent implications for the final failure attribution study in Chapter 7 are discussed.

6.4.1 Overview of the results

This study investigated how consumer knowledge differences of LCD TVs and failure cause (internal versus external cause with respect to the TV) affect attribution of product failures in the picture quality of an LCD TV. These differences were evaluated in an Internet-based experiment with 354 participants recruited via a consumer panel and various Internet forums.

The results for the control measurements demonstrated that overall the design of the Internet-based experiment was successful: there is no significant difference between perceived realism of the scenarios used, there is a significant difference in perceived picture quality between the introduction scenario and the failure scenario and approximately 80% of the participants immediately perceived the presence of a failure in the failure scenario.

Furthermore, most of the measurement scales proved to be reliable and valid. First, the results showed that the consumer knowledge measurements, including the newly developed objective expertise measurement, met the criteria for inclusion in further analyses. Despite using open response measurements of objective familiarity instead of frequency scales (which could have reduced systematic response bias (Schwarz et al., 2008, p. 27), no significant correlation is found between expertise and familiarity measurements of LCD TVs. This is similar to the results of the consumer knowledge survey discussed in Chapter 4 and further confirms the conclusions drawn in Chapter 4 and 5 that the more passive interaction with a TV in general does not automatically result in higher levels of subjective and objective expertise. The relatively high correlation (0.591) between objective and subjective expertise is in accordance with the meta-analysis results of consumer knowledge measurement correlations by Carlson et al. (2009). Secondly, the results showed that only the transformed locus scale of the causal dimension scale (Oliver, 1996; Russell, 1982) was valid and reliable in this context. Both the controllability and stability scales were therefore removed from further analyses. When reflecting on the wording of the stability and controllability items, a possible explanation for the weak validity and reliability could be that the items were too difficult to answer due to the use of failure scenarios instead of a real-life failure experience. Nevertheless, the main goal of this study was to evaluate the effect of consumer knowledge on the locus scale (see section 6.1.1). Finally, the measurements of the control variables for product involvement and product expectations (adjusted from Lancellotti (2004)) proved to be reliable and valid for further analyses. An overview of the results of the hypotheses tested in this study is shown in Table 6.9.

Most important of all, the results show that *only* the effect of objective expertise on the dependent variables is significant. Although no previous research *compared* the effect of different consumer knowledge measurements on failure attribution, this result confirms previous consumer knowledge research in which objective expertise is argued as the most reliable measurement of what consumers actually know (Brucks, 1985; Cordell, 1997) and therefore is more directly related to causal reasoning on product failures. For the analyses with objective expertise as independent variable, the effect of the control variables for product involvement and product expectations is not significant. Prior beliefs on LCD TVs and motivation to use them do not strengthen or lessen the effect of objective expertise on attribution. However, the effect of age as covariate (significant for picture quality) and failure experience as blocking factor (significant for attribution locus and picture quality) is significant.

For the analyses of the effect of consumer knowledge on correctness of the attribution (i.e. in accordance with the physical cause of the failure) of the two different failures, mixed results are found. For the noise scenario, participants with a higher level of objective expertise attribute the failure significantly more in accordance with the physical cause of the failure (i.e. external) than participants with a lower level of objective expertise. However, for the blocking artefacts scenario opposite results were found. Moreover, the results show that for attribution locus, previous experience with the failure even led to more extreme external attribution for both failure scenarios. Consequently, higher levels of objective expertise do not automatically result in attributions more in accordance with the DTV system expert's attribution of a failure.

Table 6.9 Overview of results of hypotheses testing in Chapter 6

Hypothesis	Result
H_1: Consumers with higher levels of knowledge, measured as follows: a) Objective expertise b) Subjective expertise c) Objective familiarity d) Subjective familiarity attribute product failures caused by product internal factors stronger to internal causes than do consumers with lower levels of the same measure of knowledge.	a) Rejected b) Rejected c) Rejected d) Rejected
H_2: Consumers with higher levels of knowledge, measured as follows: a) Objective expertise b) Subjective expertise c) Objective familiarity d) Subjective familiarity attribute product failures caused by product external factors stronger to external causes than do consumers with lower levels of the same measure of knowledge.	a) Accepted b) Rejected c) Rejected d) Rejected
H_3: Consumers with higher levels of knowledge, measured as follows: a) Objective expertise b) Subjective expertise c) Objective familiarity d) Subjective familiarity attribute product failures to more different causes than do consumers with lower levels of the same measure of knowledge.	a) Accepted b) Rejected c) Rejected d) Rejected
H_4: Differences in objectively measured expertise stronger relate to differences in failure attribution than differences in familiarity and subjectively measured expertise.	Accepted
H_5: Consumers with higher levels of knowledge, measured as follows: a) Objective expertise b) Subjective expertise c) Objective familiarity d) Subjective familiarity rate perceived picture quality differently than do consumers with lower levels of the same measure of knowledge.	a) Accepted b) Rejected c) Rejected d) Rejected
H_6: Consumers with higher levels of knowledge, measured as follows: a) Objective expertise b) Subjective expertise c) Objective familiarity d) Subjective familiarity rate perceived failure impact differently than do consumers with lower levels of the same measure of knowledge	a) Rejected b) Rejected c) Rejected d) Rejected

When reflecting on these results it seems that, although the blocking artefacts did not resemble artefacts caused by a digital TV signal disturbance, the participants attribute the failure in accordance with their expectations and previous experience (which led to a high percentage of attributions to the cable signal). Higher levels of objective expertise seem to further strengthen this effect. For the study discussed in Chapter 7 it is therefore interesting to investigate whether this effect is also present for other types of picture quality failures.

Next, it is found that both overall and for the individual failure scenarios, participants with higher levels of objective expertise attribute the failure to more different causes than participants with lower levels of objective expertise. Although higher levels of objective expertise result in more extreme and not necessarily more correct attributions, the attribution itself is more complex and refined.

Finally, the results show a mixed effect of objective expertise on perceived failure impact and perceived picture quality. Overall, higher levels of objective expertise (and for failure experience) result in a more negative evaluation of perceived picture quality but not in a different evaluation of perceived failure impact compared to lower levels of objective expertise. Separate analyses for each failure scenario show that the effect of objective expertise is significant for the noise scenario for both failure impact and perceived picture quality but not significant for the blocking artefacts scenario. Because these results are not consistent across the different failures and because research has shown that failure impact can also directly affect (extremity of) failure attributions (Laufer, Gillespie et al., 2005) and therefore might have influenced the results found, the separate effect of the failure impact on failure attribution will be further evaluated in Chapter 7.

6.4.2 Consumer versus designer attribution of product failures

In the study discussed in this chapter was shown that consumers attributed the two picture quality failures in LCD TVs differently than DTV system experts did. Especially for the blocking artefacts scenario, the majority of the participants attributed the failure to a “wrong” cause. This could be explained by the fact that most consumers are not familiar with the presence and/or properties of software in modern LCD TVs and therefore attributed to causes fitting with their expectations and mental model of an LCD TV. Since CE products are changing rapidly from a technological point of view, this mental model of a product’s functioning can be or can become incorrect (see for example A. Cooper (1999)). The results showed that, even despite being classified in the high objective expertise group, the participants attributed blocking artefacts to a wide range of other causes which, similar as discussed by Ceaparu et al. (2004)), can lead to ineffective problem solving strategies, higher consumer dissatisfaction and more complaints. Especially for complex products with multiple companies and service providers involved, attribution to the wrong cause and/or a wrong failure diagnosis by the manufacturer can lead to ineffective and inefficient customer service.

Therefore, analyzing (potential) product failures from an attributional point of view can contribute to a better consumer focus in the PDP by helping designers and developers to better understand and subsequently diagnose CPFs and consumer complaints. This can help to prioritize product failures from a consumer point of view and help to take the correct action for improvement of the product design.

6.4.3 Limitations and further research

The study presented in this chapter has given more insight into how consumer knowledge differences and failure cause affect failure attribution. Nevertheless, in its current form this study had several limitations. First, although the Internet-based experimental methodology enabled the use of a large sample size and the scenario-based design allowed for more control over extraneous variables, it only allowed for use of failure scenarios with highly visible failures and removed the failure from a real-life experience during product usage. Consequently, only limited inferences can be drawn to how a consumer would respond to a real failure in their own LCD TV in their home environment. Note however, that it did capture diversity in attribution of product failures and as such is a first step in investigating how consumer knowledge and failure characteristics affect failure attribution. Secondly, the classification of the open response attribution question depended on the researchers' and designers' interpretation. One could argue that other, more qualitative open response coding techniques such as content analysis, might give better insight into how consumers attribute failures.

These limitations and the potentially confounding effect of failure impact on failure attribution are addressed in the final failure attribution study discussed in Chapter 7.

7 Evaluating the effect of consumer knowledge and failure impact on failure attribution

The results of the large-scale Internet-based experiment in the previous chapter showed that the effect of consumer knowledge on failure attribution is not consistent for different types of failures. To gain deeper insight into how consumer knowledge affects failure attribution for different types of failures, this chapter uses a controlled experiment with a more homogenous group of participants and multiple failure attribution measurements. It specifically investigates how differences in consumer knowledge and how variation on the impact of a failure affect failure attribution, by conducting a laboratory experiment with an LCD TV displaying videos of failure scenarios.

This chapter is organized as follows. Section 7.1 discusses the conceptual framework and hypotheses tested in this study. Subsequently, in section 7.2 the design of the laboratory experiment to test these hypotheses is discussed. Section 7.3 reports on the results of this experiment, assessing the effect of the consumer knowledge constructs and failure impact on failure attribution. Finally, this chapter concludes with a discussion of the results and limitations of this study in section 7.4.

7.1 Conceptual framework and hypotheses

Based on the results of the Internet-based experiment discussed in the previous chapter, it was concluded that additional insight is needed into how consumers attribute product failures by using a different methodology and more in-depth attribution measurements. As discussed by Lancellotti (2004), using a multi-method approach can help to cross-validate the results and enhance the understanding of complex factors underlying failure perception.

Besides for reasons of cross-validation of the results of the study discussed in Chapter 6, the goal of the study discussed in this chapter is to investigate the possible effect of a different failure characteristic, i.e. failure impact, on failure attribution. Differences in failure impact (i.e. degree of loss of functionality) are, from a software developer's perspective, one class of criteria that decides which product development faults need to be fixed first (De Visser, 2008, p. 49). The study discussed in this chapter aims to investigate if and how differences in failure impact from a consumer point of view also affect how consumers with different levels of knowledge attribute those failures. In other words, are failures with an objectively measurable difference in failure impact also perceived differently by consumers and does this difference also affect what is perceived to be the cause of these failures? Research by Laufer, Gillespie et al. (2005) has shown for example that severity of a failure can influence the extremity of attribution. In their study, they show that in the context of product-harm crises (i.e. situations where the product is found to be defective or dangerous such as exploding car tires), observers who perceive a crisis to be more severe attribute more blame to the company than those who perceive the crisis to be less severe (Laufer, Gillespie et al., 2005). In the context of the TRADER project it is therefore interesting to investigate if and how differences in failure impact of more subtle software failures affect how consumers with different levels of knowledge attribute those failures.

To be able to compare the results with the results of the study discussed in Chapter 6, the same consumer knowledge constructs, moderating variables and dependent variables as shown in Figure 6.1 were used. To enable more in-depth investigation of differences in failure attribution, multiple failure attribution measurements and a measurement of the problem solving strategy (i.e. perceived solution to the failure shown in the failure scenario) were taken into account. For reasons of ecological validity (i.e. the degree to which the failure scenarios designed for this experiment approximate real-life failures), for this experiment only picture quality failures in an LCD TV with product internal (i.e. software) causes were used (see section 7.2.3). The conceptual research framework for this chapter is shown in Figure 7.1. Based on the above and the results of the experiment discussed in Chapter 6, five hypotheses are formulated of which four match the hypotheses formulated in Chapter 6 and hypothesis H_2 is new based on the expected effect of failure impact on failure attribution.

Hypothesis H_1: Consumers with higher levels of knowledge, measured as follows:

a) Objective expertise
b) Subjective expertise
c) Objective familiarity
d) Subjective familiarity

attribute product failures caused by product internal factors stronger to internal causes than do consumers with lower levels of the same measure of knowledge.

Hypothesis H_2: Product failures with a higher impact result in more extreme attributions than product failures with a lower impact.

Hypothesis H_3: Consumers with higher levels of knowledge, measured as follows:

a) Objective expertise
b) Subjective expertise
c) Objective familiarity
d) Subjective familiarity

attribute picture quality failures to more different causes than do consumers with lower levels of the same measure of knowledge.

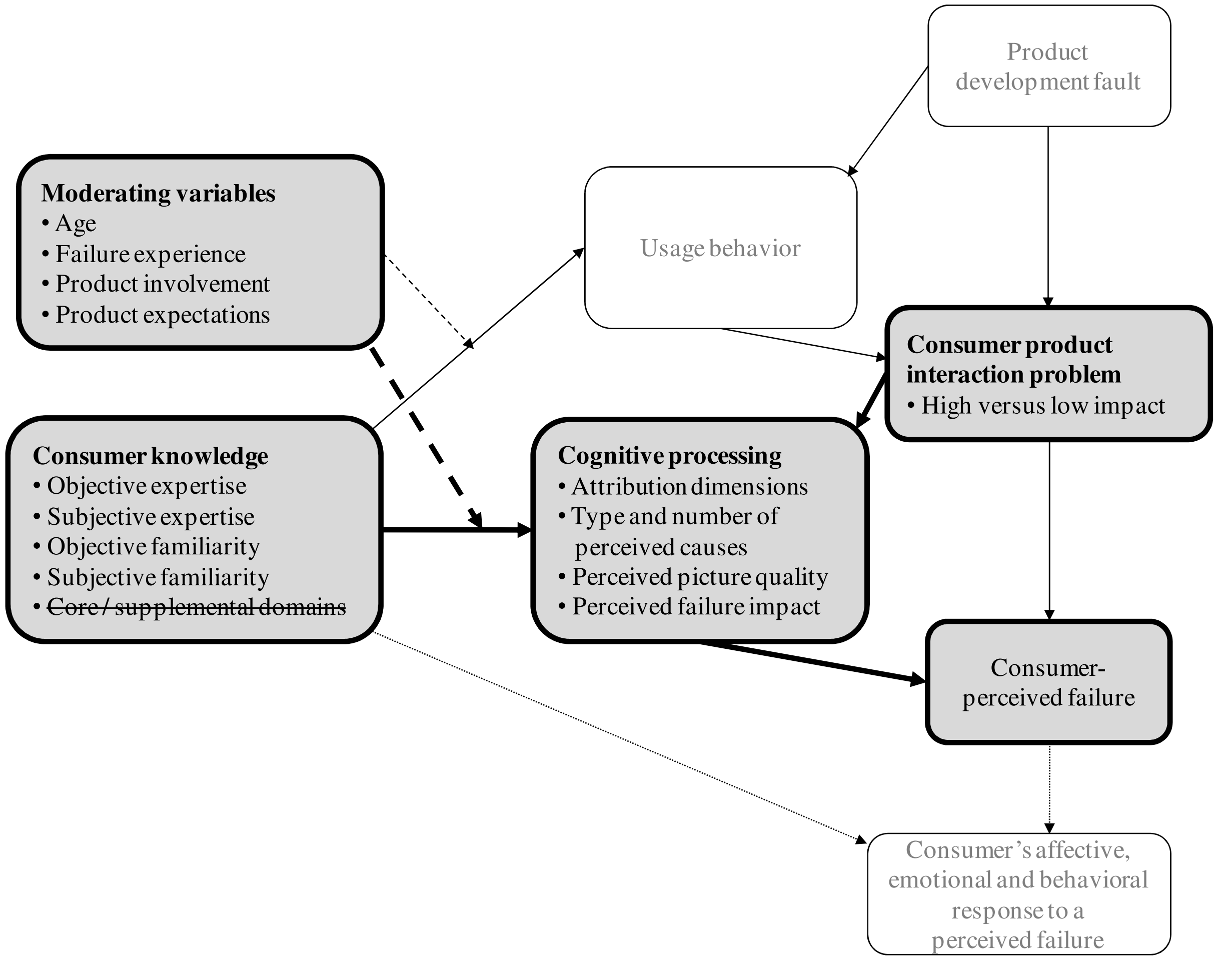


Figure 7.1 Conceptual framework to investigate the effect of consumer knowledge and failure impact on failure attribution

Hypothesis H_4: Differences in objectively measured expertise have a stronger effect on failure attribution than differences in familiarity and subjectively measured expertise.

Hypothesis H_5: Consumers with higher levels of knowledge, measured as follows:

a) Objective expertise
b) Subjective expertise
c) Objective familiarity
d) Subjective familiarity

rate perceived picture quality lower than do consumers with lower levels of the same measure of knowledge.

7.2 Method

To test the hypotheses formulated in section 7.1, a 2 (high versus low levels of consumer knowledge) x 2 (product failure with a low versus high impact) between-subjects laboratory experiment was designed in which participants were asked to evaluate a scenario showing a product failure related to picture quality of an LCD TV. This section describes the design of this experiment.

7.2.1 Experimental design

As discussed in section 6.2.1, for methodological and practical reasons a scenario-driven experiment is best suitable for investigating the effect of consumer knowledge on failure attribution for different types of picture quality failures. To enhance the real-life failure experience and to have more control over the context and setting in which the failure scenarios were evaluated, the study discussed in this chapter was conducted in a laboratory environment in which participants were shown failure scenarios on a real LCD TV. As the size of the screen on which the failure scenarios are shown could affect the visibility of failures with a low impact (i.e. loss of functionality), it was important to keep the experimental setting constant for every participant.

Furthermore, for reasons of validity it is important to select failure scenarios which significantly differ on failure impact but do not differ on the other failure characteristics (i.e. failure workaround, reproducibility, moment in use process, solvability, frequency and function importance), as defined by De Visser (2008, p. 67). To select a failure type with a significant difference in perceived degradation of picture quality within the margins of what is realistically possible in LCD TVs, a pre-test with different types of picture quality failures with varying levels of failure impact was conducted. Again for reasons of validity, the participants in this pre-test needed to be similar in characteristics to the participants in the final experiment. Consequently, the whole experiment was split up in several consecutive parts:

1. Invitation to participate in the experiment and the measurement of consumer knowledge constructs via a separate web-based questionnaire (November 2008).
2. Pre-test of different failure scenarios to select a failure type with a significant difference on failure impact (November 2008).
3. Final laboratory experiment to measure the effect of consumer knowledge and failure impact on failure attribution (December 2008).

The selection and design of the failure scenarios is further discussed in section 7.2.3.

7.2.2 Participants

As previously discussed in Chapters 3, 4 and 6, the population of interest for this research is a preferably heterogeneous population of Dutch consumers who are willing to use an LCD TV and meet generally used demographic criteria, such as aged 16 years old or above. Since it was important for this experiment to ensure that a large enough sample was achieved within the practical limitations such as resource constraints (see also the discussion of the selection of participants in section 5.3.2 for the experiment on usage behavior), a convenience sample of students and employees from various departments of Eindhoven University of Technology was used. It was ensured that none of the participants were familiar with the content of this research project. As an incentive to participate, the LCD TV used in the experiment was raffled among the participants.

Out of the 139 participants in the web-based survey (first step of the experimental design discussed in the previous section), 16 people participated in the pre-test to select the failure scenarios, three people participated in the pilot test and 58 people participated in the final experiment. The number of people who participated in the final experiment is slightly less than desired for a 2 x 2 factorial design (i.e. 4 x 20 participants per group = 80 participants (Hair et al., 2006., p. 402)). However, due to time and resource constraints and because of self selection bias towards the higher knowledge group[24], this sample size was the maximum that could be achieved and is still well above the minimum sample size required for the analyses performed (see also section 7.3.2). An overview of the demographics of the participants in the final experiment is shown in Table 7.1.

From this table can be seen that there was a bias towards younger, higher educated males, which is a consequence of using technical university students and employees as participants. Out of the 58 participants in the final experiment, all used a TV and 57 owned a TV (of which 27.6% owned an LCD and/or plasma TV and 25.9% used a digital TV signal).

Table 7.1 Overview of participant characteristics in terms of age, educational level and gender

Age	n	%	Educational level	n	%	Gender	n	%
< 21 years	6	10.3	Low	0	0.0	Male	47	81.0
21 – 30 years	43	74.2	Medium	21	36.2	Female	11	19.0
31 – 40 years	5	8.6	High	37	63.8			
41 – 50 years	3	5.2						
> 50 years	1	1.7						
Total	**58**	**100.0**		**58**	**100.0**		**58**	**100.0**

7.2.3 Selection and design of the failure scenarios

To enable the investigation of the effect of failure impact on failure attribution, first a picture quality failure in an LCD TV needs to be selected that can be differentiated on perceived failure impact. This section discusses the selection and design of different failure scenarios and the results of a pre-test to select one type of failure for the final experiment. Similar to the failure scenario selection process used in Chapter 6, input was used from DTV system experts in a brainstorm session to select two relevant, realistic failures which could be varied on

[24] i.e. Participants in the selection questionnaire with a higher level of consumer knowledge were more willing to participate in the experiment than participants with a lower level of consumer knowledge.

failure impact but were also relatively “equal” failures (in terms of the other failure characteristics). The following two failure types were chosen:

- Frame skips: fault in the software of the LCD TV or corrupted data input to the LCD TV which results in missing frames when looking at a broadcast. This can either result in a duplication of a previous frame or in a blank (i.e. black) frame (depending on the brand and type of LCD TV). Depending on the severity of the fault the frequency of missing frames can vary.
- Skin tone fault: fault in the software of the LCD TV that enhances the skin tone color shown on the display. This fault results in very bright red patches on the skin of people shown on the TV display. Depending on the type of software fault this can result in a very small disturbance on a narrow spectrum of skin tone colors to a large disturbance for a broad spectrum of skin tone colors.

Besides these two relatively unfamiliar failure types, a failure scenario with noise was included in the pre-test to form a frame of reference:

- Noise: light or severe noise on the screen due to a bad cable signal or bad weather.

For each of these failure types, two scenarios were designed: one with a light impact and one with a severe impact on picture quality (within realistic boundaries). The scenarios were designed using video editing software and evaluated by DTV system experts. For reasons of validity each failure scenario was implemented in the same fragment of a cooking program. Because picture quality evaluations are highly dependent upon the content of a video fragment (Van den Ende, De Hesselle & Meesters, 2007), it was important to select a fragment in which all failures were clearly visible and in which the content was neutral for the participants[25].

To assess which of the two failure types differed most on perceived impact, the adjectival categorical judgment method as advised in the official guidelines for the subjective assessment of TV picture quality (ITU-R Recommendation BT.500-11, 2002) was used. By using this method, random pairs of 15 second fragments of the failure scenarios were shown to the participants. After each pair, participants were asked to rate the perceived picture quality (which referred to the degree of perceived loss of functionality for picture quality failures) of the first fragment in comparison to the second fragment on a seven-point scale ranging from much worse to much better (ITU-R Recommendation BT.500-11, 2002). Each participant rated all possible 15 combinations. In this pre-test 16 people, who were selected from the 139 survey participants and who had varying levels of objective expertise on LCD TVs, participated. This pre-test took place under the same conditions and in the same context as the final experiment described in section 7.2.5 and 7.2.6.

[25] The failure scenario videos can be obtained from the author.

The scores for the paired comparisons were subsequently transformed into single overall scores for each failure scenario using multidimensional scaling in the XGms software program (Martens, 2003, Chapter 5). The resulting mean scores and error bars for a 95% confidence interval are shown in Figure 7.2. The stimuli shown on the horizontal axis in this figure can be identified as follows:

- A: Light noise
- B: Light skin tone error
- C: Light frame skips
- D: Severe noise
- E: Severe skin tone error
- F: Severe frame skips

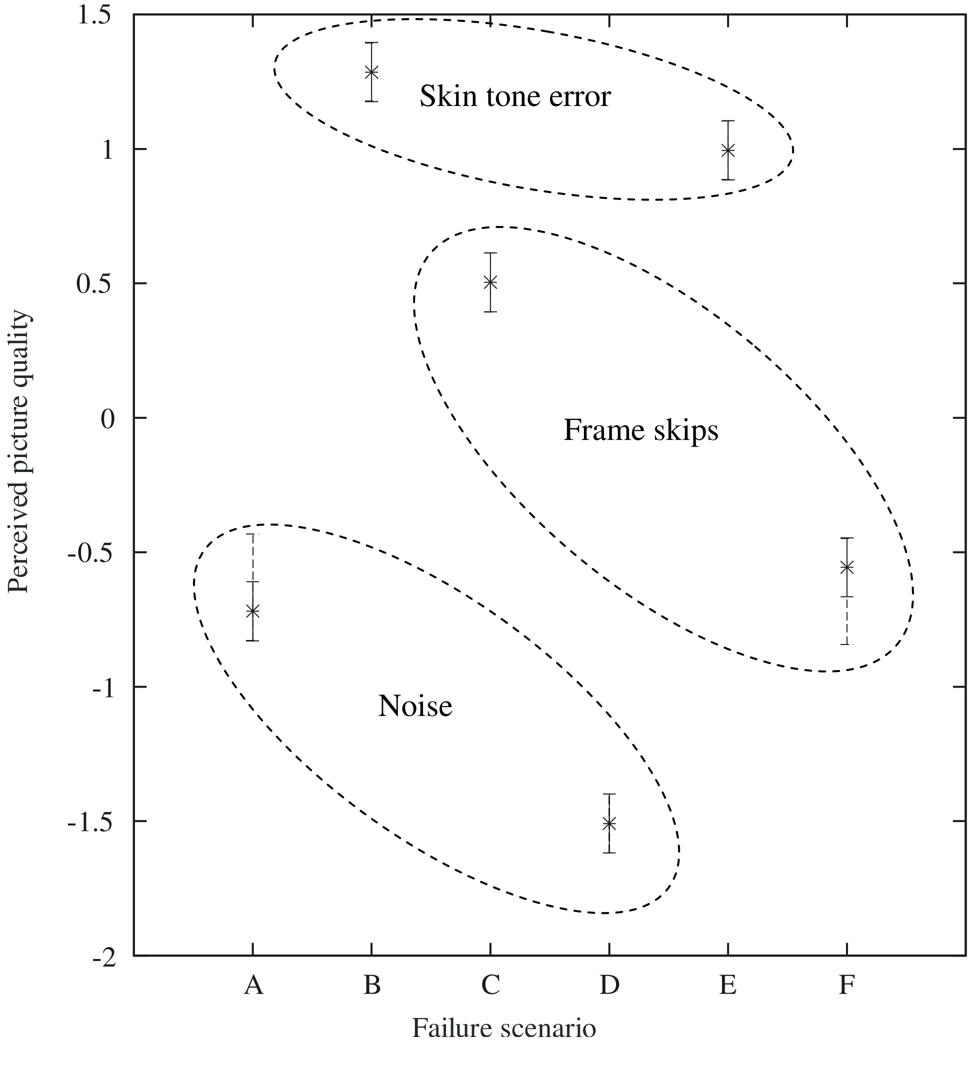


Figure 7.2 Comparison of perceived picture quality for the different failure scenarios (showing mean and error bars)

From the figure can be seen that both noise scenarios are evaluated as quite severe (negative side of perceived picture quality) while the skin tone error is overall considered the least severe failure and there is only a minimal significant difference between both skin tone error failure scenarios. Based on these results, the choice was made to use the frame skips in the final experiment. To enhance failure scenario realism and to ensure that the failure was visible for each participant, the failure scenarios for this experiment contained a one minute fragment from the same cooking program and a one minute fragment from a CNN news program.

The two minute length ensured that each failure scenario (low versus high frequency frame skips) was clearly visible and two consecutive fragments from different TV channels simulated switching channels to enhance failure scenario realism. The two failure scenarios can also be found in the CD appendix of this dissertation.

7.2.4 Experimental variables

In this section the measurements of the independent, dependent and control variables are discussed.

Consumer knowledge measurements

As shown in Figure 7.1, all the consumer knowledge constructs were used in this experiment. Because the results of the study discussed in Chapter 6 showed that all the previously used consumer knowledge measurements were valid and reliable, for the study discussed in this chapter the same measurements for subjective expertise and subjective and objective familiarity were used. These measurements can be found in Appendix 6.1.

For the objective expertise measurement, three items were added to the measurement scale previously developed and validated in Chapter 6. Although the objective expertise scale proved the be valid and reliable, because of the homogenous group of test participants it was decided to add additional items to reflect knowledge on usage of LCD TVs in failure situations more accurately (see also Brucks (1985)). The added multiple choice items can be found in Appendix 7.1.

Failure attribution measurements

To measure failure attribution several measurements were used. First of all, similar to the study discussed in Chapter 6, the adjusted and translated causal dimension scale and open response attribution measurement were used. To account for the different experimental setting in which the failure scenarios were shown to participants, the introduction text and the question formulation of the open response measurement were improved. Secondly, a "check-all-that-apply" attribution measurement was added based on the participant's answers to the failure scenarios used in Chapter 6. Finally, an open response item measuring the perceived optimal problem solving strategy was included. Both items allowed for a cross-validation with the open-response attribution measurement. All items and their response scales can be found in Appendix 7.2.

Control variables measurements
For the measurement of the control variables product involvement, product expectations and failure experience as well as for the measurements of perceived picture quality and perceived failure impact, the same measurements were used as for the study discussed in Chapter 6. These measurements can be found in Appendix 6.1.

7.2.5 Apparatus and materials

The experiment was performed in the research group's consumer test facility at the university campus. This laboratory consisted of one room in which the participants were seated on one end of a table in front of an LCD TV positioned at the other end of the table. For the experiment a 32" LCD HD ready LCD TV was used. This TV was connected to a laptop on which the videos of the failure scenarios were run. A web-based questionnaire displayed on a separate laptop was used to provide the participants instructions to perform the experiment and to record the participants' answers to the attribution and related questions. Similar to the previous experiment, for both the selection and experiment questionnaire Limesurvey (Limesurvey v1.71), run on a university campus server, was used.

7.2.6 Procedure

At the beginning of the experiment the participants were instructed that the goal of the experiment was to evaluate the quality of LCD TVs. Before starting with the experiment, the participants were provided with basic information on the LCD TV (e.g. price, time of market introduction, innovative functionalities) and the procedure of the experiment by the researcher. Subsequently, each participant was asked to read the introduction to the experiment shown in the web-based questionnaire on the laptop. Each task was explained on a separate page and the participants were asked to complete a task before proceeding with the next one. An overview of the complete experimental procedure is shown in Figure 7.3. All the items of the questionnaire and the introduction text of the failure scenario can be found in Appendix 7.2.

From Figure 7.3 can be seen that a similar ordering of questions was used as for the Internet-based experiment discussed in Chapter 6. Based on the results of the web-based experiment, the introduction to the failure scenario was improved to account for the use of video-based failure scenarios instead of a real-life failure (i.e. participants were instructed that a video of the failure was captured from an LCD TV by DTV system experts and that the capturing itself did not affect the appearance of the failure).

The experimental set-up, procedure, measurements and questionnaire were pre-tested in a pilot experiment with three participants recruited from the group of survey respondents who indicated to be willing to participate in the experiment.

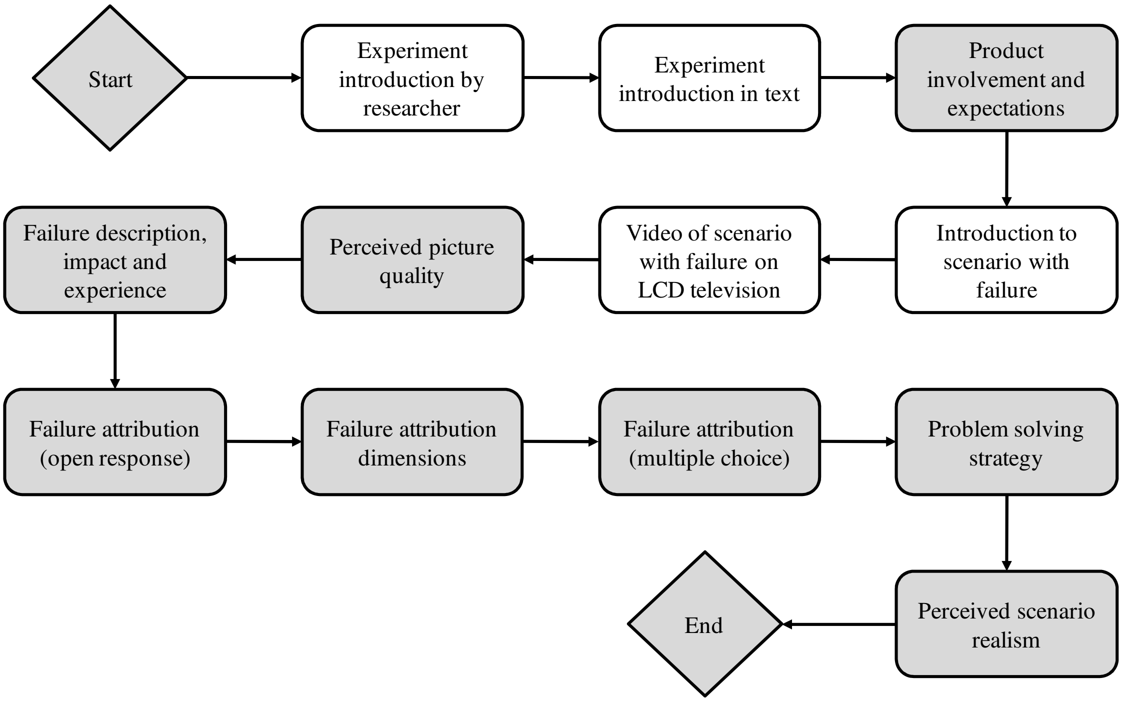


Figure 7.3 Overview of experimental procedure

7.3 Results

In this section, the results of the experiment are discussed. First, section 7.3.1 discusses the validation of the different constructs and measurement scales used throughout the experiment. Based on these results section 7.3.2 discusses the set-up of the statistical analyses and content analyses of the open response attribution measurement and problem solving strategy to test the hypotheses. Next, in section 7.3.3, the overall effect of consumer knowledge and failure impact on failure attribution is discussed. Section 7.3.4 discusses the results of the content analysis of the open response attribution measurement. Finally, section 7.3.5 discusses the results of the content analysis of the open response problem solving strategy measurement.

7.3.1 Validation of the measurements

In this section, the validation of the measurements for the independent, dependent and control variables is discussed. For all the analyses of these measurements only the experiment participants were taken into account; the selection survey only served to attract participants and did not include measurements of the control variables (which were measured *after* showing the LCD TV in the experimental setting).

To analyze the reliability and validity of the objective expertise measurement, first the scores on the individual items were calculated in the same manner as previously discussed in section 6.3.2. Subsequently, the point-biserial correlation and p-values for the individual items and Cronbach's alpha for the resulting overall scale were calculated (DIIA, 2003; Varma, n.d.).

From the results of this analysis shown in Appendix 7.3 can be seen that multiple items had to be removed from further analyses (i.e. items 2, 10 and 12 – 14). First of all, the added items which reflected different failure scenarios did not discriminate low from high knowledge participants and therefore had to be removed. Secondly, two items that were valid for differentiating consumers in the experiment discussed in Chapter 6, did not meet the criteria for inclusion in the measurement scale for the sample used in this experiment. Possibly, differences in demographics of the sample (heterogeneous versus skewed towards higher educated) could have affected the validity of these items. For the nine remaining items the Cronbach's alpha is 0.708 which is sufficiently high for further analysis (DIIA, 2003).
The results of the separate factor analyses for subjective expertise and subjective familiarity show that the items score on a single factor and measurements of Cronbach's alpha show that both scales are reliable (0.905 for subjective expertise and 0.641 for subjective familiarity).

Next, results of the factor analyses for the product involvement and expectations measurement scales show that both scales were valid. Separate measurements of Cronbach's alpha prove that both scales are (although on the lower end of the boundary) reliable (0.563 for product involvement and 0.774 for product expectations) and therefore acceptable for further analysis. The descriptive statistics for the validated consumer knowledge constructs and moderating variables are shown in Table 7.2.

Table 7.2 Descriptive statistics for the questionnaire constructs

Construct	Mean	S.D.	Scale range	Number of items
Objective expertise	6.07	2.06	0 – 9	9
Subjective expertise	3.15	1.08	1 – 5	2
Objective familiarity	12.78	7.46	2 – 35	1
Subjective familiarity	3.08	0.93	1 – 5	2
Product involvement (reverse scored)	3.82	0.63	2.33 – 5	3
Product expectations (reverse scored)	4.33	0.73	2 – 5	3

From this table can be seen that the mean objective expertise score is on the higher end of the scale, which can be attributed to the use of a convenience sample at a university.

By using the mean score of the expertise and familiarity constructs, both convergent and discriminant validity can be discussed by investigating the correlations between these constructs (see also Chapter 4 and 6). An overview of the correlations between the consumer knowledge constructs and the control variables is shown in Table 7.3. When comparing the correlations of the consumer knowledge constructs with the correlations found in the Internet-based experiment shown in Table 6.5 can be seen that these are to a large extent similar in direction and strength. However, in contrast to the results of the previous study the consumer knowledge constructs do not significantly correlate with the moderating variables. Again, this effect could be due to the use of a more highly educated sample but does not affect further analysis.

Finally, for the factor analysis of the causal dimension scale mixed results are found, similar to the results found in the study discussed in Chapter 6. The results of the factor analyses show that the stability items do not score on a single factor and were therefore removed from further analysis. Furthermore, although the controllability construct does emerge from the factor analysis, separate measurements of Cronbach's alpha (0.748 for locus and 0.429 for controllability) indicate that the scale is not reliable and therefore not suitable for further analysis. The locus scale which again was of main interest for this study proves to be both valid and reliable.

Table 7.3 Correlations (Spearman's rho) of questionnaire constructs, N = 58

	Subjective expertise	Objective familiarity	Subjective familiarity	Age	Product involvement (reversed)	Product expectations (reversed)
Objective expertise	0.591**	-0.269*	-0.134	0.043	0.067	0.214
Subjective expertise		-0.227	-0.011	0.115	0.016	0.021
Objective familiarity			0.492**	0.218	-0.096	-0.198
Subjective familiarity				0.039	-0.143	-0.109
Age					-0.258	-0.122
Product involvement						0.461**

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).

7.3.2 Set-up of the analyses

To investigate the main effect of consumer knowledge on failure attribution, multiple separate MANOVAs (for each consumer knowledge construct separately) with a 2 (consumer knowledge) x 2 (failure impact) factorial design were used in which the attribution locus scale, perceived failure impact and perceived picture quality were included as dependent variables. Participants were categorized into low versus high consumer knowledge based on a split on the mean value of the different consumer knowledge constructs. An overview of the characteristics of both groups based on a differentiation on objective expertise is shown in Table 7.4.

Results of the assumptions check for the MANOVA analysis for objective expertise showed that the data sufficiently met the criteria for performing this analysis after transformation of both the attribution locus and perceived failure impact variable. For both the MANOVA and the nonparametric tests the level of significance is set at p = 0.05. Results within the less restrictive level of p = 0.1 are indicated as marginally significant. Finally, for the MANOVA analyses the significance of the omnibus F-tests were taken from the Pillai values.

Table 7.4 Overview of participant characteristics based on a differentiation on objective expertise

	High objective expertise			**Low objective expertise**		
	Mean	**S.D.**	**Range**	**Mean**	**S.D.**	**Range**
Objective expertise	7.71	0.763	7 – 9	4.53	1.66	0 – 6
Subjective expertise	3.68	1.06	1 – 5	2.65	0.84	1 – 4
Objective familiarity	11.82	7.63	3 – 35	13.67	7.31	2 – 35
Subjective familiarity	2.95	0.98	1 – 5	3.20	0.89	2 – 5
Age (years)	25.21	6.16	20 – 52	26.57	8.34	20 – 49
Product involvement	3.88	0.56	3 – 5	3.75	0.68	2.33 – 5
Product expectations	4.49	0.62	2.67 – 5	4.19	0.81	2 – 5

For the analysis of the open response attribution and problem solving measurements, content analysis was used (Krippendorf, 1980). According to Hsieh and Shannon (2005), three major approaches to content analysis can be distinguished:

- Conventional content analysis: applicable for studies used to describe a phenomenon. The study starts with observations and codes are derived from the data itself.
- Directed content analysis: applicable for studies for which already theory or prior research exists, but this is still incomplete and needs further description. The study starts with theory or prior research and codes are derived before and during data analysis.
- Summative content analysis: applicable for studies which go beyond frequency measurements of words to discover latent variables. This study starts with keywords and these keywords are defined before and during data analysis.

Based on this overview, the type of content analysis suitable for this study is directed content analysis. Based on the results of the attribution response coding of the study discussed in Chapter 6, several categories (i.e. internal, external and mixed attributions) can already be defined and need to be further developed based on the data of the current study. For the analysis, the basic content analysis method of grouping and counting relevant attribution related words and sentences will be used (Krippendorf, 1980, p. 109).

7.3.3 Evaluation of the overall effect of consumer knowledge and failure impact

In this section, the overall effect of consumer knowledge and failure impact on the dependent variables is discussed as well as the results for the evaluation of control variables (effect of age, product involvement and product expectations) and possible confounding factors (effect of failure experience and perceived scenario realism).

Evaluation of perceived scenario realism

Before proceeding with the analysis of the main effects the design of the experiment was evaluated. A separate pair-wise Mann-Whitney U test confirmed that there is no significant difference in perceived scenario realism between the two failure scenarios ($p < 0.4$). For both scenarios the mean score of perceived scenario realism (1.71 for the low impact scenario and 2.03 for the high impact scenario) showed that both scenarios were perceived, on average, as moderately realistic (see Appendix 6.1 for the measurement scale used).

MANOVA results for the effect of consumer knowledge and failure impact

To evaluate the effect of objective expertise and failure impact on failure attribution, several MANOVA models were used as shown in the conceptual research framework in Figure 7.1. Since the effect of product involvement, product expectations and failure experience as blocking factors was not significant and did not improve the significance and power for the main independent variables, these were stepwise removed from the model (Hair et al., 2006, p. 419). Consequently, the final MANOVA model incorporated, besides the main independent variables, age as a covariate. The results of the multivariate tests and tests of between-subjects effects for this model are shown in Appendix 7.4. From this model can be seen that the overall effect of objective expertise ($F(3, 51) = 2.407$, $p < 0.1$) is only marginally significant. The results also show that the overall effect of failure impact is not significant ($F(3, 51) = 1.465$, $p < 0.3$). In other words, contrary to what was expected, the objectively differentiated and subsequently pre-tested subjective difference in failure impact did not have an overall significant effect on attribution locus, perceived picture quality and perceived failure impact. Consequently, hypothesis H_2 needs to be rejected.

Results of a separate pair-wise Mann Whitney U test further confirmed that the difference in perceived failure impact between both scenarios was not significant ($p < 0.8$). Although these differences were significant in the pre-test, due to the use of a between-subjects design and because of the experimental design, the participants lacked a frame of reference and rated both scenarios as equally severe[26].

Although the results show that the separate effect of age as covariate on the dependent variables is not significant ($F(3, 51) = 2.407$, $p < 0.1$), it did improve the significance and

---

[26] Please note that the experiment was designed as such to simulate the real-time logging of a picture quality failure and therefore no reference video with the same content could be used.

power of the main independent variables and is therefore included in the final model. Finally, no significant interactions between the dependent variables and control variables were found.

Further tests of the between-subjects effects showed that objective expertise had a significant effect on attribution locus ($F(1, 58) = 4.160$, $p < 0.05$) but not on perceived picture quality ($F(1, 58) = 1.767$, $p < 0.2$) and perceived failure impact ($F(1, 58) = 3.736$, $p < 0.4$). Based on this analysis it can be concluded that hypothesis H_1 needs to be accepted and hypothesis H_5 needs to be rejected. From the interaction between objective expertise and failure impact for attribution locus shown in Figure 7.4 can be seen that both failures were on average perceived to be caused more by TV internal factors than TV external factors (mean of attribution scale is three). For both scenarios higher levels of objective expertise also resulted in a more extreme internal attribution (i.e. attributions towards a cause inside the TV) compared with lower levels of objective expertise.

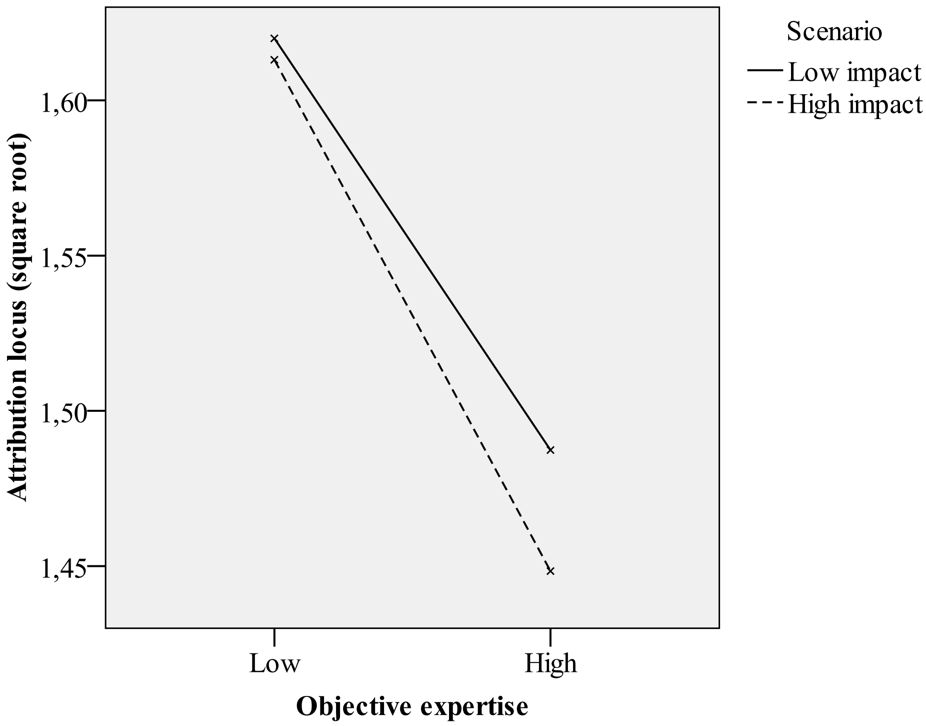


Figure 7.4 Interaction plot of objective expertise and failure impact for attribution locus (higher scores refer to more external attributions)

Furthermore, from the interaction plot between objective expertise and failure cause for perceived picture quality shown in Figure 7.5 can be seen that for both scenarios higher levels of objective expertise also resulted in a slightly more negative (although not significant) judgment of perceived picture quality compared with lower levels of objective expertise. These results are consistent with the results found in Chapter 6.

Finally, the results of similar separate MANOVAs with subjective expertise ($F(3, 51) = 0.039$, $p < 0.6$), objective familiarity ($F(3, 51) = 0.031$, $p < 0.7$) and subjective familiarity

(F (3, 51) = 0.564, p < 0.7) as main independent variable indicated that the overall effect of these consumer knowledge constructs on the dependent variables was not significant[27]. Hypothesis H_4, which stated that objective expertise stronger relates to differences in failure attribution than subjective expertise and familiarity, can therefore be accepted.

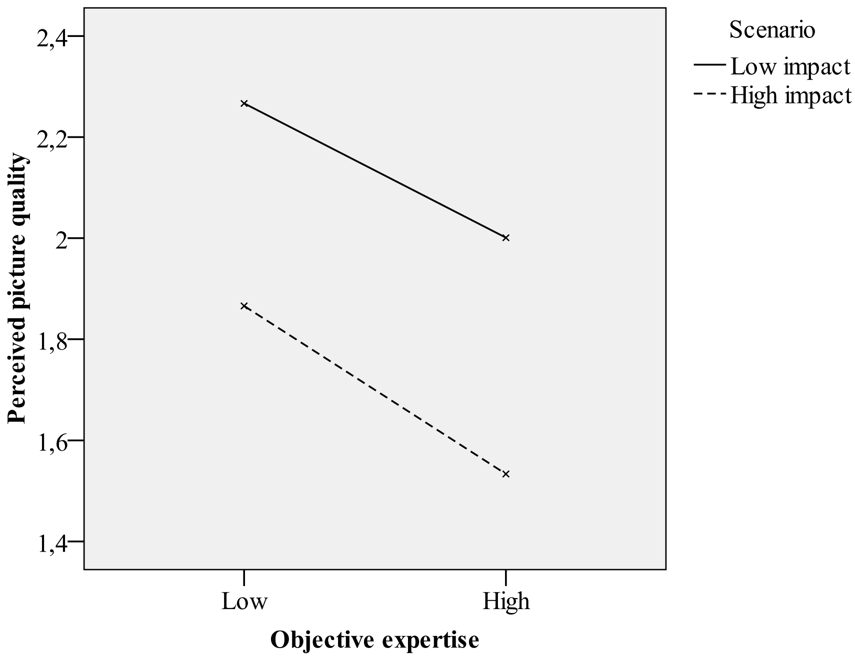


Figure 7.5 Interaction plot of objective expertise and failure impact for perceived picture quality

7.3.4 Evaluation of the open response attribution measurement

In this section, the results of the content analysis of the open response attribution measurement are discussed. The directed content analysis resulted in 23 different categories of attribution responses which could be summated into three main categories: attribution to something inside the TV, attribution to something external to the TV and attribution towards a combination of both the TV and TV signal related aspects (e.g. "frequency mismatch" refers to attribution responses in which the participant perceives that the frequency of the TV signal and the frequency of the TV display do not match which results in frame skips). The attribution to something inside to the TV is further split up into three intermediate categories. An overview of all the categories and the frequencies with which each specific category was mentioned is shown in Table 7.5. Because the differentiation on failure impact did not affect the dependent variables in the MANOVA model, for this analysis no differentiation was made between both scenarios. The results of a separate pair-wise Mann-Whitney U test show that there is no significant difference in the total number of attributed causes between participants

[27] For these analyses, similar to the model with objective expertise as independent variable, only age was taken into account as covariate.

from the low (mean = 3.23 causes) and high objective expertise group (mean = 3.32 causes) ($p < 0.9$) and therefore hypothesis H_3 needs to be rejected.

No differences between the relative frequencies of the attribution response categories for the different objective expertise groups were calculated due to the wording of the open response attribution question. This question was formulated as such that participants were asked to generate as many different plausible causes for the failure scenario as possible. Since this does not reflect confidence in one or more of the given answers, these answers did also not necessarily reflect the level of correctness of the attribution.

Table 7.5 Overview of attribution response categories and its total frequencies

Attribution response category	Total Frequency
TV_total	**83**
TV_software_total	40
• TV_software_imageprocessing	15
• TV_software_general	14
• TV_software_buffer	4
• TV_software_other	3
• TV_software_CPU	2
• TV_software_interlacing	1
• TV_software_firmware	1
TV_hardware_total	33
• TV_hardware_general	7
• TV_hardware_powersupply	7
• TV_hardware_wire	6
• TV_hardware_backlight	5
• TV_hardware_crystals	5
• TV_hardware_screen	3
TV_general	10
External_total	**89**
• External_coaxdefect	16
• External_signalquality	15
• External_coaxconnection	14
• External_provider	13
• Exernal_powersupply	9
• External_interference	9
• External_settings	7
• External_signalsplitter	6
Combination_total	**18**
• Combination_signal-TVmismatch	10
• Combination_frequencymismatch	8

Separate analyses of the multiple-choice attribution measurement did also not show any significant differences in selection of the response categories between the low and high objective expertise groups. However, the content analysis of the open response attribution measurement did result in valuable information from a designer's perspective by showing a very large and diverse spectrum of attributed causes and is therefore an essential step in attribution response analysis. The implications of these results are further discussed in section 7.4.

7.3.5 Evaluation of the open response measurement of problem solving strategies

To conclude the results section, the content analysis of the open response measurement of the problem solving strategy is discussed. The directed content analysis resulted in 22 different categories of attribution responses which could again be summated into three main categories: solve the problem indirectly by looking for help (e.g. by consulting the helpdesk or reading the manual), solve the problem directly with a strategy related to the functioning of the TV and solve the problem directly with a strategy related to the outside of the TV (e.g. checking the quality of the TV signal or inserting the power plug into another power socket). An overview of the categories and the frequencies with which each specific category was mentioned is shown in Table 7.6.

Similar to the results of the content analysis of the open response attribution measurement, the results shown in Table 7.6 demonstrate that participants have a large spectrum of problem solving strategies. Based on a ranking of frequencies, three main strategies emerge from the data: take the TV back to the shop or manufacturer, switch the TV "off" and back "on" (which resets the software) and check another cable for similar interference (in other words, verify whether the coax cable is the source of the problem). According to attribution of the frame skips scenarios by DTV system experts, the latter two strategies most likely directly result in a successful solution to the problem. However, both strategies only refer to less than half of the total number of mentioned problem solving strategies. Many strategies mentioned do not (directly) result in a solution to the problem or imply contact is needed with the shop or manufacturer. The further implications of this study for the design of complex CE and on the insight into the effect of consumer diversity on CPFs, is discussed in the following section.

Table 7.6 Overview of problem solving strategy categories and its total frequencies

Problem solving strategy category	Total frequency
External_total	**32**
External_shop	17
External_Internet	8
External_manual	2
External_expert	2
External_cable provider	2
External_helpdesk	1
Inside TV_total	**41**
Inside TV_switch TV off	21
Inside TV_switch TV channel	8
Inside TV_change TV settings	4
Inside TV_reset TV	3
Inside TV_reset TV settings	3
Inside TV_let TV cool off	1
Inside TV_upgrade firmware	1
Outside TV_total	**57**
Outside TV_check other cable	26
Outside TV_try other TV	8
Outside TV_remove cable	6
Outside TV_check other input	5
Outside TV_check signal quality	5
Outside TV_change power socket	3
Outside TV_check other appliances	2
Outside TV_check meter cupboard	1
Outside TV_ask neighbours	1

7.4 Conclusion and discussion

This study investigated how consumer knowledge differences of LCD TVs and failure impact (low versus high impact) affect attribution of product failures in the picture quality of an LCD TV. These differences were evaluated in a laboratory experiment with 58 participants recruited at the university.

First of all, most of the measurement scales proved to be valid and reliable in a similar degree to the results discussed in Chapter 6. The results did show that the added objective expertise items could not discriminate low from high knowledge consumers and were therefore not further included in the scale. On top of that, two other previously validated items in Chapter 6 proved to be either too easy or lacked discriminatory power for the narrower sample used in the study discussed in this chapter. This shows that the formation of an objective expertise measurement is, besides the goal for which the measurement is used, also dependent upon the population considered in the research project. Correlations between the consumer knowledge constructs proved to be very similar to those already discussed in section 6.4.1.

Furthermore, again similar to the results of Chapter 6, the results of this study showed that the attribution stability and controllability measurements scales are not valid and reliable in the experimental set-up used in this chapter. Although only the validated locus scales is of importance for this research, further research on the consequences of differences in attributions needs to take this into account.

Next, the results of the hypotheses tested in this chapter demonstrated similarities but also differences with the conclusions of Chapter 6. An overview of the results of the hypotheses tested in this study is shown in Table 7.7.

Table 7.7 Overview of results of hypotheses testing in chapter 7

Hypothesis	Result
H_1: Consumers with higher levels of knowledge, measured as follows: a) Objective expertise b) Subjective expertise c) Objective familiarity d) Subjective familiarity attribute product failures caused by product internal factors stronger to internal causes than do consumers with lower levels of the same measure of knowledge.	a) Accepted b) Rejected c) Rejected d) Rejected
H_2: Product failures with a higher impact result in more extreme attributions than product failures with a lower impact.	Rejected
H_3: Consumers with higher levels of knowledge, measured as follows: a) Objective expertise b) Subjective expertise c) Objective familiarity d) Subjective familiarity attribute product failures to more different causes than do consumers with lower levels of the same measure of knowledge.	a) Rejected b) Rejected c) Rejected d) Rejected
H_4: Differences in objectively measured expertise stronger relate to differences in failure attribution than differences in familiarity and subjectively measured expertise.	Accepted
H_5: Consumers with higher levels of knowledge, measured as follows: a) Objective expertise b) Subjective expertise c) Objective familiarity d) Subjective familiarity rate perceived picture quality lower than consumers with lower levels of product related knowledge.	a) Rejected b) Rejected c) Rejected d) Rejected

One of the most striking results of this study is the lack of any significant effect of the variation of failure impact on failure attribution. Although the specially designed pre-test to select a relevant and significantly differentiable failure on failure impact demonstrates a significant difference in perceived picture quality for the two frame skips scenarios, the results of the experiment do not reflect these results. A possible explanation for this result could be that the differentiation on failure impact is not large enough for a sufficient effect failure attribution in an experiment with an acceptable but not very large sample. For studies which do show a significant effect of failure impact on failure attribution, very large differences in failure impact were used (e.g. a scratch on a product versus situations in which consumers were physically harmed) (Silvera & Laufer, 2005). However, from both a methodological and practical perspective this was not relevant in the context of the research presented in this dissertation since this research concerns more subtle software related and realistic failures and is limited to one picture quality failure (in order to avoid the potentially confounding effect of other failure characteristics on failure attribution). Since for both scenarios the frame skips were clearly visible in a two minute video fragment, this apparently triggered similar attribution responses.

From the results of the hypotheses testing shown in Table 7.7 can also be seen that, similar to the results of the study discussed in Chapter 6, only the effect of objective expertise was significant (albeit marginally). This further strengthens the conclusion drawn in Chapter 6 that the objective expertise measurement is the most reliable measurement of what consumers actually know and therefore more accurately reflects differences in failure attribution than other consumer knowledge constructs.

The results of the analysis of the attribution locus scale showed that when the participants were forced to make a choice between internal and external attribution of the frame skips scenario, for both scenarios the participants from the high objective expertise group attributed the failure more in accordance with the most likely physical cause of the failure. When comparing this with the results of Chapter 6, it again shows that the effect of objective expertise on attribution correctness depends on the type of failure and previous experience with related (but not the same) failures with comparable effects on picture quality. In contrast to the results of Chapter 6, the differences between the lower and higher levels of objective expertise for the number of attributed causes and perceived picture quality are not significant. A possible explanation for this difference could be the relatively small spread on the objective expertise measurement for the sample used in this study.

Although the adjusted open response attribution and problem solving strategy measurements were not suitable for analyzing differences between the consumer knowledge groups, the results of the content analysis give designers insight into the large spectrum of perceived causes and perceived strategies to solve the problems. Further implications of these results are discussed in the overall conclusions in the following chapter.

Overall, the study presented in this chapter has given more insight into how consumer knowledge differences and failure impact affect failure attribution and how failure attribution can be qualitatively measured and analyzed. This chapter concludes the empirical research studies presented in this Ph.D. dissertation. In the following final chapter overall conclusions are drawn and the theoretical and practical implications of this dissertation are discussed.

8 Conclusions and Discussion

Previous research showed that existing approaches for managing product Quality and Reliability (Q&R) do not cover uncertainties associated with the increase of consumer complaints for complex Consumer Electronics (CE). Part of the problem is that there is lack of consumer insight regarding the relation between the diversity of consumers and the propagation of product development faults to Consumer-Perceived Failures (CPFs) and consumer complaints. To support effective consumer-focused decision making in product development, for example with respect to failure prioritization from a consumer perspective, this dissertation aims to provide more insight into this relation. This chapter concludes this dissertation by discussing the key findings and contributions of the research presented in the previous chapters.

This chapter is organized as follows. Section 8.1 gives an overview of the key research findings of this dissertation. In section 8.2 the contributions of this research are discussed for both theory and practice. Subsequently, in section 8.3 the generalization of the research findings is discussed. In section 8.4 the limitations of the conducted research are discussed. Finally, in section 8.5 recommendations for future research are discussed.

8.1 Summary of key findings

This section summarizes the key findings of this dissertation. First, in section 8.1.1 an overview is given of the research context, problem and research questions addressed in this dissertation. Subsequently, in sections 8.1.2, 8.1.3 and 8.1.4 the research findings for each of the sub research questions are discussed. Finally, in section 8.1.5 overall conclusions are drawn.

8.1.1 Research overview

Chapter 1 discussed that the lack of consumer insight with respect to the propagation of product development faults to consumer complaints can be attributed to two factors. In the field of CE there are on the one hand increasingly demanding, more fragmented and less predictable consumer groups, while on the other hand product designers and developers find it more difficult to predict consumer dissatisfaction and complaints for these diverse groups. To address the lack of understanding of the relation between product development faults and consumer complaints in the classical Q&R fault-complaint propagation model, insights from consumer behavior literature were used to develop a fault-complaint propagation model from a consumer perspective in Chapter 2. This model incorporates Q&R problems from a consumer perspective: *Consumer-Perceived Failures*. Because currently used consumer segmentation criteria do not sufficiently cover differences in CPFs, this dissertation validated a part of the fault-complaint propagation model by investigating the effect of differences on multiple dimensions of a single consumer characteristic which affects the consumers' understanding of complex CE: "consumer knowledge". In the context of this research, consumer knowledge relates to both the cognitive structures consumers have of a product's functioning and the cognitive processes to be able to perform product-related tasks successfully (Alba & Hutchinson, 1987).

Based on the results of an explorative experiment with an implemented failure in the teletext functionality of a TV and on insights from HCI and consumer behavior research, a conceptual framework was developed in Chapter 3 to better understand the underlying factors of the propagation of product development faults to CPFs and its relation with consumer knowledge. Both product usage behavior and failure attribution were identified as important consumer-dependent mediating variables of this propagation. This conceptual framework is shown in Figure 8.1, followed by an overview of the conducted empirical research in Table 8.1.

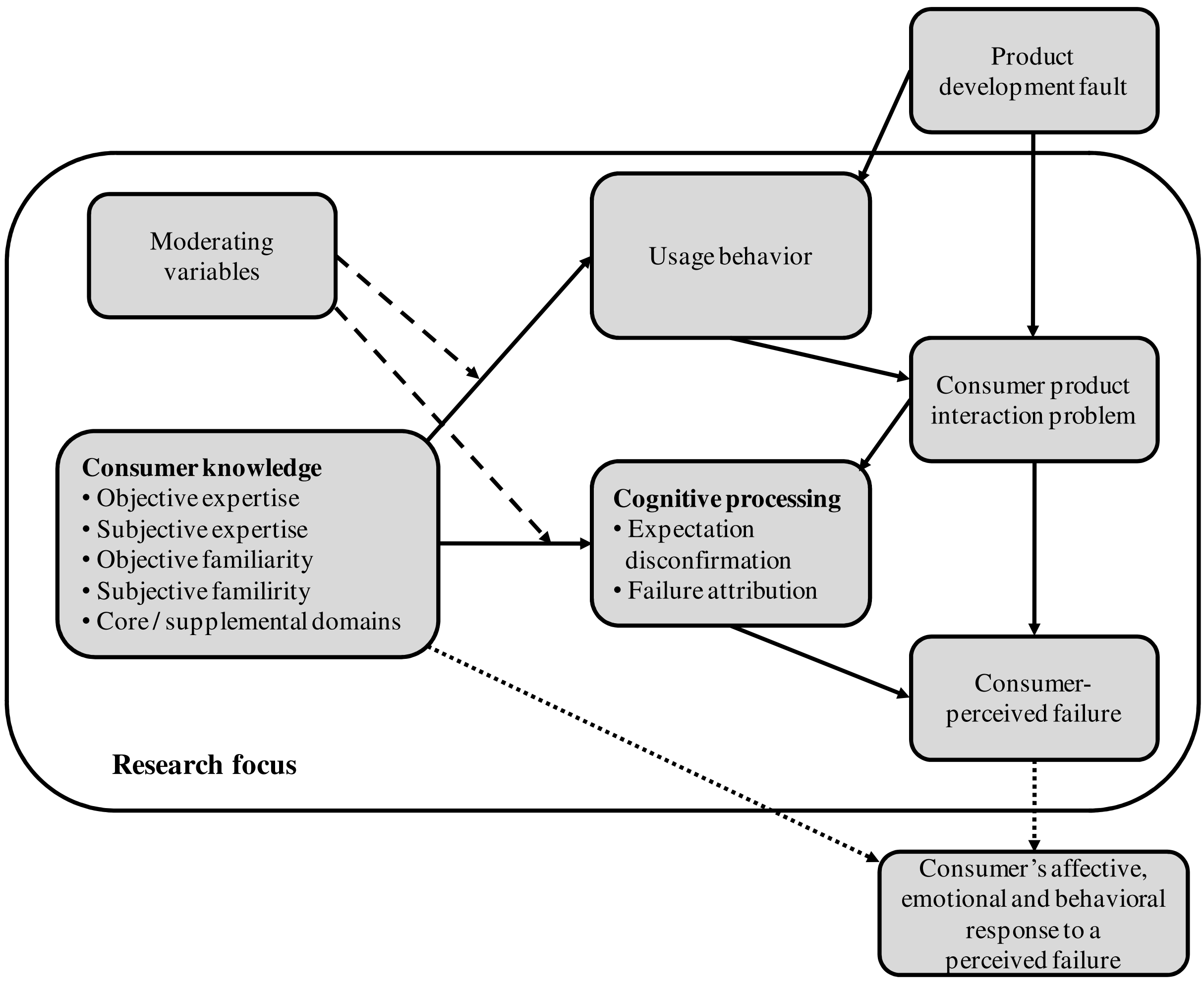


Figure 8.1 Overview of the conceptual framework used in this dissertation

Table 8.1 Overview of set-up of empirical research

Chapter	4	5	6	7
Research question(s) addressed	Differentiation of consumers on subjective expertise, and objective and subjective familiarity on the core and supplemental knowledge domains of CE	Investigation of the effect of consumer knowledge on product usage behavior	Differentiation of consumer on objective expertise of CE and investigation of the effect of consumer knowledge and failure cause on failure attribution	Investigation of the effect of consumer knowledge and failure impact on failure attribution
Methodology	Paper-based survey	Laboratory quasi-experiment	Web-based quasi-experiments	Laboratory quasi-experiment
Consumer knowledge constructs	• Subjective expertise Based on validated scale Flynn and Goldmith (1999) • Subjective familiarity Newly developed scale based on validated constructs • Objective familiarity Newly developed scale based on validated scale in different contexts • Core and supplemental domains Newly selected domains based on comparable research	See Chapter 4	• Objective expertise Newly developed scale based on constructs from Brucks (1985), Arning and Ziefle (2008) and Cordell (1997) • Subjective expertise Shortened scale based on results of Chapter 4 • Objective familiarity Adjusted scale based on results Chapter 4 • Subjective familiarity Shortened scale based on results Chapter 4	• Objective expertise Adjusted based on results Chapter 6 • Subjective expertise Same as Chapter 6 • Objective familiarity Same as Chapter 6 • Subjective familiarity Same as Chapter 6
Dependent variables	None	• Effectiveness Based on validated measurements Hornbæk (2006) and ISO 9241-11 (1998) • Efficiency Based on validated measurements from Hornbæk (2006) and ISO 9241-11 (1998) • Usage patterns Based on validated process mining measurements (Van der Aalst et al., 2007)	• Attribution dimensions Adjusted scale from Russell (1982; 1987) • Number and type of attributed causes Newly developed scale • Perceived picture quality Newly developed scale • Perceived failure impact Adjusted scale from De Visser (2008)	• Attribution dimensions Adjusted based on results Chapter 6 • Number and type of attributed causes Adjusted based on results Chapter 6 • Perceived picture quality Same as Chapter 6 • Perceived failure impact Same as Chapter 6

Based on this conceptual framework, the following main research question was addressed in the remainder of the dissertation:

How does consumer knowledge affect usage behavior and failure attribution of consumer electronics?

To answer this research question, three sub research questions were derived:

1. *How can consumers be differentiated on knowledge of consumer electronics?*
2. *How does consumer knowledge affect usage behavior of consumer electronics?*
3. *How does consumer knowledge affect attribution of product failures in consumer electronics?*

An overview of the main conclusions for each of the sub research questions is shown on the next page in Table 8.2. In the following sections the findings for each of these sub research questions are discussed in more detail, followed by a section in which the overall conclusions for the main research question are discussed.

8.1.2 Differentiation of consumers on knowledge of complex CE

Across all the studies conducted in Chapter 4, 6 and 7, several conclusions can be drawn on the differentiation of consumers on knowledge of complex CE. First of all, the research presented in this dissertation demonstrated the successful development and validation of measurements for different consumer knowledge constructs of (LCD) TVs. The final measurement scales for subjective expertise (two items), subjective familiarity (three items) and objective familiarity (one item) used in Chapter 6 and 7 proved to have sufficient convergent and discriminant validity and acceptable levels of scale reliability (Cronbach's alpha = 0.6 – 0.8). For objective expertise a completely new, product specific, scale (11 multiple choice and check-all-that-apply items) was developed based upon a literature review and input from DTV system experts. This scale proved to be reliable (Cronbach's alpha > 0.8) and all the items demonstrated acceptable levels of item difficulty and discriminatory power for a heterogeneous sample (Chapter 6). For more homogenous samples, such as university students and employees (Chapter 7), some items lacked discriminatory power and/or were too easy and were removed. Consequently, the complete objective expertise scale is only valid for measuring factual knowledge of LCD TVs for heterogeneous samples and caution must be taken when applying this measurement in another context or for a different sample.

Table 8.2 Overview of main conclusions for the sub research questions

Research question	How can consumers be differentiated on knowledge of consumer electronics?	How does consumer knowledge affect usage behavior of consumer electronics?	How does consumer knowledge affect attribution of product failures in consumer electronics?
Main conclusions	• Successful development and validation of measurements of: a) Subjective expertise (two items, based on validated scale) b) Objective expertise (11 items, new scale) c) Subjective familiarity (three items, new scale) d) Objective familiarity (1 item, new scale) • Mixed validity for the use of knowledge domains to differentiate consumers: only objective measurements can be used when differentiating consumers on multiple domains of CE • Overall similar correlations between consumer knowledge constructs compared to previous research, apart from correlations between objective familiarity of TVs and objective / subjective expertise of TVs • Age, gender and educational level affect level of consumer knowledge	• Significant effect of subjective expertise on product usage behavior • Significant effect of objective familiarity of computers on usage behavior while the effect of objective familiarity of TVs is not significant • Stronger effect of subjective expertise on usage behavior than the overall effect of familiarity • Participants with lower levels of subjective expertise experience more and different interaction problems than participants with higher levels of subjective expertise	• Out of the investigated effect of subjective / objective familiarity and expertise constructs, only differences in objective expertise significantly affect failure attribution • Higher levels of objective expertise result in: 1) more extreme (i.e. convinced) attributions; 2) more refined attributions with more dimensions but 3) not necessarily more correct attributions. • Both failure cause and failure impact do not significantly affect how consumers attribute failures

Secondly, the results of the study discussed in Chapter 4 showed mixed results for the validity of the use of core and supplemental domains to differentiate consumers on knowledge of complex CE. The results showed that the subjective expertise and subjective familiarity scales of both knowledge domains are significantly correlated (0.6 and 0.5 respectively) which resulted in an unequal distribution of consumers across the four hypothesized consumer knowledge segments. Although the items refer to distinctly different product domains, from a consumer perspective the *perceived* level of expertise and *perceived* level of familiarity appear to refer to a more general level of confidence in using and interest in CE respectively. Because a differentiation on objective familiarity of both knowledge domains resulted in an equal distribution of consumers among the consumer knowledge segments, the results of this study suggest that only objective measurements of consumer knowledge can be used when differentiating consumers on multiple knowledge domains of CE.

Next, across the studies similar correlations between consumer knowledge constructs were observed. In the studies discussed in Chapters 6 and 7, a significant and similar correlation of 0.591 was observed between objective and subjective expertise of LCD TVs which is in line with the results of a meta-analysis on consumer knowledge construct correlations by Carlson et al. (2009). However, in contrast to previous research on consumer knowledge measurements of complex CE (e.g. Cordell (1997), the results of the three different surveys show that there is no correlation between objective familiarity of TVs and subjective or objective expertise of TVs. As, for the computer domain, these constructs did significantly correlate in line with previous research, these results suggest that usage experience of TVs, which is mostly a relatively “passive” form of interaction (compared to digital cameras or computers), not necessarily results in higher levels of (perceived) expertise. Consequences of these differences are further discussed in section 8.2.

Finally, when comparing the differentiation of consumers on consumer knowledge across the different studies, the results clearly show an effect of the heterogeneity of the sample on the spread of the consumer knowledge measurements. Although the separate confounding effect of demographic variables on the dependent variables was limited, the results show that age, gender and educational level (Chapter 7) affect the level of consumer knowledge for different measurements (e.g. age positively affects the level of usage experience and negatively affects objective and subjective expertise of TVs). Consequently, the selection of consumer knowledge constructs as criterion for differentiating consumers for a consumer test depends on the target consumer group for a product (e.g. a very narrow homogeneous consumer group versus mass consumer markets), the type of product (e.g. passive versus active interaction) and the goal of the consumer test.

8.1.3 Effect of consumer knowledge on product usage behavior

The results of the laboratory experiment discussed in Chapter 5 show that participants with higher levels of subjective expertise are able to complete significantly more tasks, need significantly less time and detour steps to complete the tasks and displayed usage patterns which conform significantly more to the designers' usage model than participants with lower levels of subjective expertise. For the effect of subjective expertise on the number of steps needed to complete the task and for the interaction effects between subjective expertise and task complexity no significant results were found. Due to the high correlation between the subjective expertise measurements, no differences between the effect of subjective expertise on the core and supplemental domains on product usage behavior were found.

However, for the effect of objective familiarity, compared to the hypothesis, opposite results were found: the overall effect of objective familiarity of computers on task effectiveness and efficiency is significant while the effect of objective familiarity of TVs is not significant. This opposite result could be explained by the fact that usage experience of TVs mostly relates to a relatively "passive" form of interaction while usage experience of computers can relate to a broader scope of tasks and interactions. Consequently, only partial support was found for explaining differences in product usage behavior by differentiating on core and supplemental knowledge domains. Since subjective expertise mostly relates to confidence in the level of knowledge on a certain product, the effect of subjective expertise on two different but also similar technically complex products does not differ that much. Nevertheless, the overall effect of subjective expertise on product usage behavior is significantly stronger than the overall effect of objective familiarity. In other words, differences in *perceived* factual knowledge stronger relate to differences in consumer behavior than differences in usage experience. Although the effect of experience with a product or task (i.e. objective familiarity) on usability has already been validated in usability literature (e.g. see Nielsen (1993) and Ziefle (2002)), the findings of this experiment demonstrate that there is an even stronger effect for expertise constructs.

Finally, the results of the process mining analysis showed that for the dual screen and switch channel task, participants with lower subjective expertise experienced more and different interaction problems than participants with higher levels of subjective expertise. The inconsistent effect for the digital picture task could be explained by the experimental design and/or the "plug & play" design of the multimedia browser interface. Participants with a higher level of subjective expertise first needed to understand the automatic start-up of this functionality while participants with a lower level of expertise automatically followed the support given by the information on the UI. These findings help product developers and designers to better understand differences in product usage behavior when consumers encounter interaction problems and they can therefore help to take better design decisions.

8.1.4 Effect of consumer knowledge on failure attribution

To investigate the effect of consumer knowledge on failure attribution, an Internet-based and a laboratory experiment were conducted. Given the methodological and resource constraints of this research project, a choice was made to use a scenario-based evaluation of the perception of picture quality failures in LCD TVs while differentiating on two different failure characteristics: failure cause (i.e. a failure cause by TV internal versus TV external factors) and failure impact (i.e. a failure with high versus low impact on TV picture quality).

First of all, the results of both studies confirmed the hypothesis stating that objective expertise has a stronger effect on failure attribution than the other consumer knowledge constructs. Moreover, *only* the effect of objective expertise on failure attribution is significant. This result shows that only objective expertise differences affect differences in consumer perception of product failures. This has important implications because currently used test methods often differentiate consumers on previous experience (i.e. familiarity) with a product.

Secondly, the results of both studies demonstrate that both failure cause and failure impact do not significantly affect how consumers attribute the failures. Although previous research by Silvera and Laufer (2005) shows that extreme differentiation on failure impact affects the extremity of attributions, the results of the study in Chapter 7 show that within the practical limitations for realism of failure scenarios in CE, this effect is not significant. However, in both studies the effect of objective expertise on attribution locus is significant, albeit not consistently in accordance with the physical cause of the failure. For both the noise scenario (Chapter 6) and the frame skips scenario (Chapter 7), consumers with higher levels of objective expertise attribute the failure more in accordance with the physical cause of the failure. In contrast, for the blocking artefacts scenario (Chapter 6), consumers with higher levels of objective expertise attribute the failure stronger to external causes, which is not in accordance with the physical cause of the failure. Consequently, the results of both studies show that higher levels of objective expertise on a product do not automatically result in attributions that are more in accordance with the real physical cause of the failure. It seems that the effect of objective expertise on attribution locus depends on both the type of failure and previous experience with related (but not the same) failures with a comparable effect on the functioning of a product (i.e. blocking artefacts are common when using a digital cable signal but usually have a different duration and impact on picture quality compared to the blocking artefacts caused by software faults).

The results of the coding of the open response attribution measurements in both studies show that consumers, when asked to reflect on all possible realistic causes of the failures shown in the scenarios, attribute the failure to a large spectrum of causes of which many are considered as highly unlikely by DTV system experts. Furthermore, the results of the Internet-based experiment support the hypothesis that consumers with higher levels of objective expertise attribute failures to more different causes than consumers with lower levels of objective expertise. Although higher levels of objective expertise result in more extreme and not necessarily more correct attributions, the attribution itself has more dimensions and is more refined. The analysis of the results of the study discussed in Chapter 7 do not show a similar significant effect which could be due to the small spread on objective expertise for the sample used in this study.

Finally, the failure attribution studies show that both age (for both studies) and failure experience (Chapter 6) affect the strength of the effect of objective expertise on failure attribution. The results also show that prior beliefs on LCD TVs and motivation to use an LCD TV do not strengthen or lessen the effect of objective expertise on failure attribution. Consequently, these results confirm that the effect of consumer knowledge on consumer behavior must be studied with specifically tailored moderating variables to evaluate the strength of the effect.

8.1.5 General conclusions

This dissertation started with the observation that the field of CE is increasingly challenging for product design and development. One of those challenges is that product designers and developers need more insight into differences in consumer groups beyond using themselves as "target user" (A. Cooper, 1999, p. 17; Norman, 1998, p. 155; Hasdoğan, 1996) to be able to better predict and prevent consumer dissatisfaction and complaints (Den Ouden, 2006, p. 58). This research addresses this gap and shows that it is too easy to simply state that the consumers and designers perceive a product's functioning and failures differently.

This research shows that, when evaluating the effect of consumer diversity on fault-complaint propagation, consumer knowledge can be used to differentiate product use and failure attribution for DTV systems. However, especially for failure attribution this effect is not consistent across different types of failures and is in most cases only significant for objective expertise differences that are not commonly addressed in consumer profiles. Especially in the context of fast evolving complex CE, objective expertise measurements are important because familiarity or subjective expertise measurements on the (technical) functioning of currently available products can quickly become "incorrect" or "incomplete" for the next generation of products.

This research also demonstrates that the classical Q&R fault-complaint propagation model does not cover all potential reasons for consumer dissatisfaction and complaints. The adjusted fault-complaint propagation model from a consumer perspective, which incorporates the consumer's *perception* of product failures, incorporates a broader spectrum of potential antecedents and provides insight into how consumer diversity can affect this propagation on different levels and via different consumer behavioral mechanisms.

In the increasingly challenging field of CE, differentiation on consumer knowledge constructs and the applied attributional approach with which differences in the consumer's perception and reasoning about product failures and its causes can be evaluated, can potentially give product designers and developers insight into how design decisions affect the occurrence of CPFs and consumer dissatisfaction for diverse consumer groups early in PDPs.

8.2 Research contributions

8.2.1 Theoretical implications

Based on the conclusions drawn in section 8.1, the following implications for theory can be deducted.

Insight into fault-complaint propagation from a consumer perspective
The first important theoretical contribution of this research project is the modeling of the propagation of product development faults to consumer complaints by combining aspects from both classical Q&R and consumer behavioral models. This model captures more potential sources of consumer dissatisfaction and complaints (i.e. product development faults, consumer diversity and the usage environment) than classical Q&R models and can therefore potentially help to better understand the underlying factors of the propagation of product development faults to consumer complaints in practice.

Set-up and validation of consumer knowledge measurements for complex CE
In this dissertation, measurements of four different consumer knowledge constructs (i.e. subjective and objective familiarity and expertise) of LCD TVs (and partly for computers) were either newly developed or selected and adjusted from previous research. The data collected by using three different surveys with different samples support the validity and reliability of these measurement scales and they can therefore be applied in further research on LCD TVs and for other CE (in the case of subjective expertise, subjective familiarity and objective familiarity).

Furthermore, the analyses of the correlations between the consumer knowledge constructs contribute to consumer behavior research by showing that, in contrast with previous research on for example digital cameras (Cordell, 1997), not for all product categories an increase in product usage experience in terms of frequency in duration of use automatically results in higher levels of product expertise. For complex CE, adjusted measurements of deeper level interactions (e.g. using advanced menus) beyond the usage of basic functionalities are needed as indicator for relevant usage experience which potentially leads to an increase in product expertise.

Attribution of *technological* product failures
A third theoretical contribution of this dissertation is the validation of attribution measurements in product-related failure situations of CE and the insight into the effect of consumer knowledge on failure attribution. Although attribution theory itself is well-founded in consumer behavior research and many studies addressed the attribution of *service* failures, few papers could be found which specifically addressed how consumers arrive at attribution of *product* failures (Folkes, 1988; Silvera & Laufer, 2005; Weiner, 2000).

The findings of this dissertation contribute to attribution theory by specifically showing how failure attribution can be measured for technological product failures and by showing how different consumer knowledge constructs affect open-response and locus measurements of this attribution. These results therefore further validate the results discussed in the failure attribution study on photo development failures by Somasundaram (1993). Finally, the results of this dissertation also show that more research is needed on the applicability and measurement of the controllability (i.e. who is perceived to be in control of the cause of the failure) and stability dimensions (i.e. is the failure perceived as being stable over time or erratic) in the context of attribution of technological product failures. Since all the attribution dimensions *together* can be a predictor of consumer complaint behavior and expectations on the type of redress after submitting a complaint, it is important to further investigate the applicability and measurement of these dimensions for failures in complex CE.

Contribution to validation of the UPFS research model developed by De Visser (2008)
A final theoretical contribution of this research is the partial validation of the UPFS research model developed by De Visser (2008). Although this dissertation did not focus on the level of product failure impact assessment, the results did provide insight into the effect of several consumer characteristics (i.e. consumer knowledge and age) on CPFs which could eventually result in differences in perceived failure severity. As such, the combination of the results of both research projects contributes to a better understanding of which factors affect consumer dissatisfaction and consumer complaints and supports failure prioritization and design decision making from a consumer perspective in the PDP of complex CE.

8.2.2 Practical implications

Besides theoretical contributions, the following contributions to practice can be deducted from the results of this research project.

Accounting for consumer diversity in product use and failure evaluation
First of all, the results presented in this dissertation have practical implications for the selection of consumers for consumer tests when investigating differences in product use and failure evaluation. While in practice participants for consumer and usability tests are predominantly selected based on demographic and lifestyle based product adoption models including a differentiation on product usage experience, the results of this research show that differences on deeper level expertise measurements more strongly reflect differences in product use and failure attribution than differences in usage experience. Consequently, when specifically investigating whether consumers understand a product's functioning, taking expertise differences into account better reflects differences in consumer evaluation of products and product failures. Depending on the specific goal of the study, the results of this dissertation can help to select the appropriate consumer knowledge constructs to differentiate consumer groups. For example, for investigating differences in usage behavior the easy-to-apply subjective expertise or familiarity measurements can suffice while for investigating

differences in *perception* of failure causes tailored objective expertise measurements are required.

Attributional approach to gain insight into consumer perception of product failures

For the investigation of the effect of consumer knowledge on failure attribution, an attributional approach was developed in Chapter 6 and 7 which can also be applied in practice to evaluate consumer perception of product failures. The used approach can contribute to consumer test practice on two aspects:

- <u>The use of a scenario-based approach to evaluate consumer perception of product failures</u>: in this dissertation was shown how to set-up and validate failure scenarios to evaluate how consumers perceive product failures. Although care must be taken when projecting results of a scenario-based approach on real-life usage situations, it can help designers to gain insight into how consumers deal with potential product failures in the early stages of a PDP.
- <u>Validated failure attribution measurements</u>: in this dissertation failure attribution measurements were adjusted to and validated in the context of CE products. These measurements can be easily applied in practice when evaluating how consumers perceive and respond to product failures.

Improving feedback loops in the PDP

The insights presented in this dissertation can be used to help customer service centers and product designers and developers to better understand and subsequently diagnose CPFs and consumer complaints, for field feedback as well as for consumer tests. This can help to prioritize product failures from a consumer point of view and help to take the correct action for improvement of the product design. Moreover, when also taking the other attribution dimensions (i.e. perceived stability and controllability of the failure) into account, failure attribution is even a predictor for consumer complaint behavior and type of redress expected (Folkes, 1984).

Supporting consumers during use of complex CE

Understanding when and how consumers attribute product failures can give insight to customer call centers to give better support to consumers when they experience problems. It can also help designers to change design aspects to influence attribution. For example, in the UI of several high-end TVs, on-screen information is added to guide the consumer when the TV detects a bad signal quality. Furthermore, such insights can be used to improve the manual and “help” instructions. Given the increasing number of complaints (Den Ouden, 2006) and because it cannot be assumed that technology always works perfectly, research on how to support consumers during use of complex technological products is essential when aiming for a better consumer experience.

Implications for TRADER

Finally, the results of this research project contribute to the TRADER project by showing the complexity of factors underlying a consumer's response to software reliability problems in DTV systems. Not only the effect of software reliability improvements on aspects of system reliability but also the *consumer's perception* of these aspects of system reliability should be taken into account. For a reliable TV of the future from a consumer's perspective, failures need to be analyzed, minimized and prioritized from *both* software reliability and a consumer's perception point of view.

8.3 Generalization

The research presented in this dissertation specifically dealt with investigating the impact of different levels of consumer knowledge on product use and failure evaluation for innovative LCD televisions. When generalizing this research in the context of increasingly ambient intelligent products, it shows that methods such as "think like a consumer" or "expert reviews" are insufficient in helping product designers and developers to cover the large spectrum of problems related to how different consumers groups use these products and perceive their behavior because:

- Classic mental models of products do not match with new ambient intelligent products. The mental models of product designers and developers overlap with a product's architecture but new ambient intelligent products are highly dynamic and are used by highly dynamic consumer groups. These dynamic consumers groups will have varying levels of (subjective and objective) knowledge on different knowledge domains related to the product which will increasingly create potential for a mismatch of their mental model with the mental model of the product designers and developers.
- Products no longer function in an isolated environment but operate together with a class of products and services in a network together with the people who use it.

To address these developments, product designers and developers need to be aware that consumer diversity can no longer be addressed only at the end of a PDP, but needs to be addressed as part of an adaptive loop of design and use of a product.

Finally, it is important to note that because this research project and the TRADER project combined insights from multiple disciplines (i.e. Technology Management, Marketing, Information Systems and Consumer Behavior), this ensured that the research findings are embedded in a relevant research and practical context. Although multidisciplinary research presents challenges from both a methodological and a research presentation (e.g. wherever one wants to present the research findings, they have to be rewritten to fit in a certain scientific discipline) point of view, it does provide more insight into product development and product usage in practice beyond a mono-disciplinary consumer behavioral or Q&R study.

8.4 Limitations

As with all research, the present study contained several limitations of which the most important are discussed in this section.

First of all, to account for the diversity of consumers and consumer groups in real life, a quasi-experimental methodology with convenience samples was used in this dissertation. As such, consumers were not randomly assigned to groups but were assigned based on occurring characteristics. Because of this decision, only the strength and direction of the relationship between consumer knowledge and product use and failure evaluation is validated, but no conclusive causal relations can be drawn (Goodwin, 2005, p. 315; Stangor, 1998, Chapter 9). Furthermore, for both the experiment discussed in Chapter 5 and the experiment discussed in Chapter 7, the number of participants was less than preferred for MANOVA analyses (due to time and resource constraints and because participants had to meet several inclusion criteria) but still meet the minimum required sample size to allow for the use of such an analysis. Therefore, the results need to be carefully evaluated and more research with a larger sample size is needed to further validate the findings of these experiments.

Secondly, since all surveys and experiments were conducted by using a single product type, i.e. an LCD TV, care needs to be taken when generalizing these results for complex CE in general. However, analogies can be drawn with other CE where the consumer experience with the media content on the device is an important determinant of consumer satisfaction, where usage experience (in terms of frequency and duration of use) not necessarily also refers to higher levels of expertise and where the product's functioning (and perception of its functioning) depends on its interactions with other products and services. Consider for example a smart phone that can be used to watch online YouTube videos for which the quality of the consumer experience depends on multiple parties (e.g. the device manufacturer, the software developer, the service provider, the media content provider etc.). Next, as discussed by De Visser (2008, p. 153), for other complex CE, similarly structured PDPs, business trends, complex technologies and diversity in consumer groups apply. Based on these similarities one can argue that the findings of this dissertation are also potentially applicable for other complex CE such as multimedia entertainment centers or wireless music stations.

Next, for the evaluation of the effect of consumer knowledge on failure attribution, mostly video-based scenarios of implemented failures in picture quality of LCD TVs were used. Although the selection and design of the failure scenarios was extensively pre-tested and reviewed by DTV system experts and the experimental designs were successfully validated, care must be taken when generalizing these results for consumer complaints in real-life product usage situations. For example, experiencing a failure in a product you have recently bought probably leads to a more negative response than when experiencing a failure in a simulated failure setting for a product you do not own. Despite these disadvantages, the use of scenarios allowed for the exploration of more different types of failures in more controllable settings than if they had to be realistically reproduced in a real LCD TV in a laboratory setting.

Finally, in Chapter 6 a web-based experiment with embedded video-based failure scenarios was used to evaluate the spectrum of, and differences in, the consumer's attribution of different types of product failures. Although great care was taken to ensure the validity and reliability of this method, the use of this method does also not allow for generalization of the results of this experiment to the consumer's attribution of the same failure in real-life product usage situations.

8.5 Recommendations for future research

The research findings and conclusions discussed in the previous sections were fully validated within the defined boundaries of this research project. Besides these contributions, the research findings also provide suggestions for new research directions of which the most important recommendations are discussed in this section.

Future research further validating the conceptual fault-complaint propagation model
The validation of the complete fault-complaint propagation model requires, similar as for the validation of the UPFS model developed by De Visser (2008), a gradual approach in which the (combined) effect of different antecedents (consumer characteristics, usage environment and product development faults) on the different parts of the propagation model is further evaluated.

Furthermore, for the studies discussed in this dissertation mainly video-based scenarios of picture quality failures in LCD TVs were used. Although these failure scenarios were carefully selected and designed in accordance with input from DTV system experts, future research should investigate the effect of consumer knowledge on product usage and failure evaluation for other products and failure types using different types of failure scenarios. For example, future research could explore the possibilities to implement product failures in realistic usage situations in which the participants had previous experience with the product before being asked to reason about failure causes.

Finally, for both practical and methodological reasons the study discussed in Chapter 5 did not evaluate the effect of objective expertise on usage behavior. Future research should investigate this effect to further increase the generalization of differences between the effects of the different consumer knowledge constructs.

Guiding product use and failure evaluation through design
The findings reported in the dissertation also call upon further research to investigate whether changes in the product design itself can influence or reduce the effect of consumer knowledge on product usage behavior (e.g. in case of "plug & play" functionalities) and whether changes

in the product design can be used to influence attribution (e.g. adding on-screen display information on the TV). An interesting research question in this context is whether specific cues in the UI, when consumers for example experience degradation in TV picture quality, can “guide” attributions and as such can lead to more effective and efficient problem solving strategies and a higher level of satisfaction.

Knowledge miscalibration

Finally, an interesting direction for future research could be to evaluate the separate effect of so-called “knowledge miscalibration” (i.e. scoring high on subjective expertise and low on objective expertise or the other way around) (Alba & Hutchinson, 2000; Carlson et al., 2009) on product use and failure attribution. Although research on this phenomenon is still limited, investigating this effect could be interesting in the context of fast evolving technology in CE where a high level of objective expertise on a current product can quickly become outdated for a future product. The result could be a miscalibration of knowledge for discontinuous and even incremental innovations which in itself could have a stronger effect on decision making in product failure situations than the level of factual knowledge itself (Carlson et al., 2009).

References

Aarts, E. & Ecarnação, J.L. (Eds.) (2006). *True visions – The emergence of ambient intelligence*. Berlin: Springer-Verlag.

Alba, J.W. & Hutchinson, J.W. (1987). Dimensions of consumer expertise. *Journal of Consumer Research, 13(4),* 411–454.

Alba, J.W. & Hutchinson, J.W. (2000). Knowledge calibration: What consumers know and what they think they know. *The Journal of Consumer Research, 27(2)*, 123–156.

Arning, K. & Ziefle, M. (2008). Development and validation of a computer expertise questionnaire for older adults. *Behaviour & Information Technology, 27(1)*, 89–93.

Arning, K. & Ziefle, M. (2009). Effects of age, cognitive and personal factors on PDA menu navigation performance. *Behaviour & Information Technology, 28(3)*, 251–268.

Aviezinis, A., Laprie, J.-C., Randell, B. & Landwehr, C. (2004). Basic concepts and taxonomy of dependable and secure computing. *IEEE Transaction on Dependable and Secure Computing, 1(1)*, 11–33.

Baron, R.M. & Kenny, D.A. (1986). The moderator – mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51(6)*, 1173–1182.

Baskoro, G., Rouvroye, J.L., Brombacher, A.C. & Radford, N. (2003). High contrast consumer test: a case study. *In proceedings of the European Safety and Reliability Conference 2003*, (pp. 107–111).

Battarbee, K. (2004). Co-experience – understanding user experiences in social interaction. Doctoral dissertation, University of Art and Design, Finland.

Behling, O. & Law, K.S. (2000). *Translating questionnaires and other research instruments – Problems and solutions.* Series: Quantitative applications in the social sciences. London: Sage Publications Inc.

Bekker, M. & Long, J. (2000). User involvement in the design of human-computer interactions: Some similarities and differences between design approaches, In S. McDonald, Y. Waern & G. Cockton (Eds.), *People and Computers XIV- Usability or Else!*, (pp. 135–148). Springer.

Belkin, N.J. (2000). Helping people find what they don't know. *Communications of the ACM, 43(8)*, 5–61.

Berden, T.P.J., Brombacher, A.C. & Sander, P.C. (2000). The building bricks of product quality: An overview of some basic concepts and principles. *International Journal of Production Economics, 67*, 3–15.

Berkman, A.E. & Erbuğ, ç. (2005). Accommodating individual differences in usability studies on consumer products. *In Proceedings of the 11th Conference on Human Computer Interaction Vol. 3*.

Breedveld, K., Van den Broek, A., De Haan, J. Harms, L., Huysmans, F. & Van Ingen, E. (2006). *De tijd als Spiegel – Hoe Nederlanders hun tijd besteden*. Den Haag: Sociaal Cultureel Planbureau, October 2006.

Boersma, J., Loke, G., Loh, H.T., Lu, Y. & Brombacher, A.C. (2003). Reducing product rejection via a High Contrast Consumer Test. *In Proceedings of the European Safety and Reliability Conference 2003*, (pp. 191–193).

Broadbridge, A. & Marshall, J. (1995). Consumer complaint behaviour: The case of electrical goods. *International Journal of Retail & Distribution Management, 23(9)*, 8–18.

Brombacher, A.C., Sander, P.C., Sonnemans, P.J.M. & Rouvroye, J.L. (2005). Managing product reliability in business processes ‘under pressure’. *Reliability Engineering and System Safety, 88*, 137–146.

Brucks, M. (1985). The effects of product class knowledge on information search behavior. *Journal of Consumer Research, 12(1)*, 1–16.

Carlson, J.P., Vincent, L.H., Hardesty, D.M. & Bearden, W.O. (2009). Objective and subjective knowledge relationships: A quantitative analysis of consumer research findings. *Journal of Consumer Research, 35,* 864–876.

Ceaparu, I., Lazar, J., Bessiere, K., Robinson, J. & Schneiderman, B. (2004). Determining causes and severity of end-user frustration. *International Journal of Human-Computer Interaction , 17(3)*, 333–356.

Chillarge, R. (1996). What is software failure? *IEEE Transactions on Reliability, 45(3)*, 354–355.

Christiaans, H.C.M., Fraaij, A.L.A., De Graaff, E. & Hendriks, C.F. (2004). *Methodologie van technisch-wetenschappelijk onderzoek*. Utrecht: Uitgeverij Lemma BV.

Cook, T.D. & Campbell, D.T. (1979). *Quasi-experimentation – Design & analysis issues for field settings*. Chicago: Rand McNally College Publishing Company.

Cooper, A. (1999). *The inmates are running the asylum: Why high-tech products drive us crazy and how to restore the sanity*. Indianapolis: Sams publishing.

Cooper, R.G. (1999). From experience: The invisible success factors in product innovation. *Journal of Product Innovation Management, 16*, pp. 115–133.

Cooper, R.G. (2001). *Winning at new products: Accelerating the process from idea to launch*. New York: Perseus Publishing.

Cooper, R.G. (2005). New products – What separates the winners from the losers and what drives success. In K.B. Kahn, G. Castellion, & A. Griffin (Eds.), *The PDMA Handbook of New Product Development* (2nd ed., pp. 279–301). New Jersey: John Wiley and Sons, Inc.

Cordell, V.V. (1997). Consumer knowledge measures as predictors in product evaluation. *Psychology & Marketing, 14(3)*, 241–260.

Cuomo, D.L. (1994). Understanding the applicability of sequential data analysis techniques for analysing usability data. *Behaviour & Information Technology, 13(1)*, 171–182.

Darnell, M.J. (2008). Making digital TV easier for less-technically-inclined people. In Masthoff, S. Panabaker, M. Sullivan, A. Lugmayr (Eds.), *1st International Conference on Designing Interactive User Experiences for TV and Video: ACM International Conference Proceeding Series, Vol. 291*, (pp. 27–30). New York: ACM Press.

Davis, F.D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13(3)*, 319–340.

Day, R.L. & Landon, E.L. Jr. (1977). Toward a theory of consumer complaining behavior. In A.G. Woodside, J.N. Sheth & P.P. Bennet (Eds.), *Consumer and industrial buying behavior* (pp. 425–437). New York: Elsevier North Holland.

De Leeuw, E.D. (2008). Choosing the method of data collection. In E.D. de Leeuw, Hox, J.J. & Dillman, D.A. (Eds.), *International Handbook of Survey Methodology* (pp. 113–135). New York: Lawrence Erlbaum Associates.

De Leeuw, E.D. & Hox, J.J. (2008). Self-administered questionnaires: Mail surveys and other applications. In E.D. de Leeuw, Hox, J.J. & Dillman, D.A. (Eds.), *International Handbook of Survey Methodology* (pp. 239–263). New York: Lawrence Erlbaum Associates.

De Leeuw, E.D., Hox, J.J. & Dillman, D.A. (2008). The cornerstones of survey research. In E.D. de Leeuw, Hox, J.J. & Dillman, D.A. (Eds.), *International Handbook of Survey Methodology* (pp. 1–17). New York: Lawrence Erlbaum Associates.

De Marez, L.S.B. & Verleye, G.B.M. (2004). ICT-innovations today: making traditional diffusion patterns obsolete, and preliminary insight of increased importance. *Telematics and Informatics, 21*, 235–260.

De Visser, I.M. (2008). *Analyzing user-perceived failure severity in consumer electronics products: Incorporating the user perspective into the development process*. Doctoral dissertation, Eindhoven University of Technology, The Netherlands.

Den Ouden, E. (2006). *Development of a design analysis model for consumer complaints*. Doctoral dissertation, Eindhoven University of Technology, The Netherlands.

Desmet, P. & Hekkert, P. (2007). Framework of product experience. *International Journal of Design, 1(1)*, 57–66.

DIIA (2003). *Test item analysis & decision making*. Division of Instructional Innovation and Assessment, University of Texas at Austin, USA.

Dillman, D.A. (2000). *Mail and internet surveys: The tailored design method*, second edition. New York: John Wiley & Sons, Inc.

Dillon, A. & Watson, C. (1996). User analysis in the HCI – the historical lessons from individual differences research. *International Journal of Human-Computer Studies, 45*, 619–637.

Docampo Rama, M. (2001). *Technology generations handling complex user interfaces*. Doctoral dissertation, Eindhoven University of Technology, The Netherlands.

Dodd, T.H., Laverie, D.A., Wilcox, J.F. & Duhan, J.F. (2005). Differential effects of experience, subjective knowledge, and objective knowledge on sources of information used in wine purchasing. *Journal of Hospitality & Tourism Research, 29(1)*, 3–19.

Engel, J.F., Blackwell, R.D. & Miniard, P.W. (1995). *Consumer behaviour*. London: Dryden Press.

Feng, J. & Sears, A. (2009). Beyond errors: measuring reliability for error-prone interaction devices. *Behaviour & Information Technology 2009 Ifirst article*, 1–15.

Fischer, W. (2004). *Digital television: A practical guide for engineers*. Berlin: Springer.

Flynn, L.R. & Goldsmith, R.E. (1999). A short, reliable measure of subjective knowledge. *Journal of Business Research, 46*, 57–66.

Folkes, V.S. (1984). Consumer reactions to product failure: An attributional approach. *Journal of Consumer Research, 10(4)*, 398–401.

Folkes, V.S. (1988). Recent attribution research in consumer behaviour: A review and new directions. *The Journal of Consumer Research, 14(4)*, 548–565.

Fournier, S. & Mick, D.G. (1999). Rediscovering satisfaction. *Journal of Marketing, 63(4)*, 5–23.

Fowler, F.J. Jr. & Cosenza, C. (2008). Writing effective questions. In E.D. de Leeuw, Hox, J.J. & Dillman, D.A. (Eds.), *International Handbook of Survey Methodology* (pp. 136–160). New York: Lawrence Erlbaum Associates.

Goodwin, C.J. (2005). *Research in psychology – Methods and design*. Crawfordsville: John Wiley & Sons Inc.

Gould, J.D. & Lewis, C. (1985). Designing for usability: Key principles and what designers think. *Communications of the ACM, 28(3)*, 300–311.

Griffin, A. & Hauser, J.R. (1993). The voice of the customer. *Marketing Science, 12(1)*, 1–27.

Grudin, J. (1991). Systematic sources of suboptimal interface design in large product development organizations. *Human-Computer Interaction, 6*, 147–196.

Grudin, J. & Pruitt, J. (2002). Personas, participatory design and product development: An infrastructure for engagement. *In Proceedings Participatory Design conference 2002*, (pp. 144–161).

Günther, C.W. & Van der Aalst, W.M.P. (2006). A generic import framework for process event logs. In J. Eder, S. Dustdar (Eds.), Business Process Management Workshops. *In Proceedings BPM 2006 International Workshops*, Vienna, Austria, September 4-7, 2006. *Lecture Notes in Computer Science, 4103*, pp 81–92. Berlin: Springer.

Hair, J.F. Jr., Black, W.C., Babin, B.J., Anderson, R.E. & Tatham, R.L. (2006). *Multivariate data analysis*. Sixth Edition. New Jersey: Pearson - Prentice Hall.

Han, S.H., Yun, M.H., Kwahk, J. & Hong, S.W. (2001). Usability of consumer electronic products. *International Journal of Industrial Ergonomics, 28*, 143–151.

Harris, K.E., Mohr, L.A. & Bernhardt, K.L. (2006). Online service failure, consumer attributions and expectations. *Journal of Services Marketing, 20(7)*, 453–458.

Hasdoğan, G. (1996). The role of user models in product design for assessment of user needs. *Design Studies, 17*, 19–33.

Heider, F. (1958). *Psychology of interpersonal relations*. USA: John Wiley & Sons Inc.

Herstatt, C. & Von Hippel, E. (1992). From experience: Developing new product concepts via the lead user method - a case study in a “low-tech” field. *Journal of Product Innovation Management, 9*, 213–221.

Hilbert, D.M. & Redmiles, D.F. (2000). Extracting usability information from user interface events. *ACM Computing Surveys, 32(4)*, 384–421.

Hornbæk, K. (2006). Current practice in measuring usability: Challenges to usability studies and research. *International Journal of Human-Computer Studies, 64*, 79–102.

Hornix, P. (2007). *Performance analysis of business processes through process mining*. Master’s thesis, Eindhoven University of Technology, The Netherlands.

Horrigan, J.B. (2008, November 16). When technology fails. *Pew Research Center Publications – Internet Project Data Memo*. Retrieved March 19, 2009, from http://www.pewinternet.org/~/media//Files/Reports/2008/PIP_Tech_Failure.pdf.pdf.

Howell, D.C. (2002). *Statistical methods for psychology*, 5th edition. Pacific Grove: Thompson Learning.

Hsieh, H.-F. & Shannon, S.E. (2005). Three approach to qualitative content analysis. *Qualitative Health Research, 15(9)*, 1277–1288.

ISO 9241-11 (1998). *ISO 9241-11: Guidance on usability*. International Organization for Standardization.

ITU-R Recommendation BT.500-11 (2002). Methodology for the subjective assessment of the quality of television pictures. *International Telecommunication Union*, Geneva, Switzerland.

Jones, E.E. & Davis, K.E. (1965). From acts to dispositions: The attribution process in person perception. *Advances in Experimental Social Psychology, 2*, 219–266.

Johnson, E.J. & Russo, J. E (1984). Product familiarity and learning new information. *The Journal of Consumer Research, 11(1)*, 542–550.

Kanis, H. (1998). Usage centred research for everyday product design. *Applied Ergonomics, 29(1)*, 75–82.

Karapanos, E., Zimmerman, J., Forlizzi, J. & Martens, J.-B. (2009). User experience of time: An initial framework. In *Proceedings of the 27th International Conference on Human Factors in Computing Systems*, (pp. 729–738). Boston: ACM Press.

Kaulio, M.A. (1998). Customer, consumer and user involvement in product development: A framework and a review of selected methods. *Total Quality Management, 9(1)*, 141–149.

Kelley, H.H. (1967). Attribution theory in social psychology. In *Proceedings Nebraska Symposium on Motivation 15*, (pp. 192 – 238). Lincoln: University of Nebraska Press.

Kelley, H.H. & Michela, L. (1980). Attribution theory and research. *Annual Review of Psychology, 31*, 457–501.

Ketola, P. (2002). *Integrating usability with concurrent engineering in mobile phone development*. Doctoral dissertation, University of Tampere, Finland.

Kieras, D.E. & Bovair, S. (1984). The role of a mental model in learning how to operate a device. *Cognitive Science, 8*, 255–273.

Kim, J. & Han, S.H. (2008). A methodology for developing a usability index of consumer electronic products. *International Journal of Industrial Ergonomics, 38*, 333–345.

Koca, A. & Brombacher, A.C. (2008). Extracting "broken expectations" from call center records: Why and how. In *Extended Abstracts on Human Factors in Computing Systems*, (pp. 2985–2990). Florence: ACM Press.

Koca, A., Funk, M., Karapanos, E., Rozinat, A., Van der Aalst, W.M.P., Corporaal, H., Martens, J.B.O.S., Van der Putten, P.H.A., Weijters, A.J.M.M. & Brombacher, A.C. (2009). Soft reliability: An interdisciplinary approach with a user-system focus. *Quality and Reliability Engineering International, 25(3)*, 3–20.

Koca, A., Karapanos, E. & Brombacher, A.C. (2009). 'Broken expectations' from a global business perspective. In *Proceedings of the 27th International Conference Extended Abstracts on Human Factors in Computing Systems*, (pp. 4267–4272). Boston: ACM Press.

Kowalski, R.M. (1996). Complaints and complaining: Functions, antecedents and consequences. *Psychological Bulletin, 119(2)*, 179–196.

Krazit, T. (2008, August 12). Analyst: Infineon chipset possible cause of Iphone 3G issues. *CNET News.* Retrieved March 19, 2009, from http://news.cnet.com/8301-13579_3-10015301-37.html?part=rss&subj=news&tag=2547-1040_3-0-5.

Krippendorf, K. (1980). *Content analysis: An introduction to its methodology*. London: SAGE Publications.

Kujala, S. & Mäntylä, M. (2000). Studying users for developing usable and useful products. *In Proceedings of 1st Nordic Conference on Computer-Human Interaction*, (pp. 1–11).

Kujala, S. & Kauppinen, M. (2006). Identifying and selecting users for user-centered design. In *Proceedings of the Third Nordic Conference on Human-Computer Interaction*, (pp. 297–303).

Lancellotti, M.P. (2004). *Technological product failure: The consumer coping process*. Doctoral dissertation, University of Southern California, United States of America.

Laprie, J.C. (1985). Dependable computing and fault tolerance: Concepts and terminology. *In Proc. 15th IEEE International Symposium on Fault-Tolerant Computing (FTCS-15)*, (pp. 2-11).

Laufer, D., Gillespie, K., McBride, B. & Gonzalez, S. (2005). The role of severity in consumer attributions of blame: Defensive attributions in product-harm crises in mexico. *Journal of International Consumer Marketing, 17 (2/3)*, 33–50.

Laufer, D., Silvera, D.H. & Meyer, T. (2005). Exploring differences between older and younger consumers in attributions of blame for product harm crises. *Academy of Marketing Science Review 9(7).*

Lazar, J. & Norcio, A. (2003). Training novice users in developing strategies for responding to errors when browsing the web. *International Journal of Human-Computer Interaction, 15(3)*, 361–377.

Lazar, J., Meiselwitz, G. & Norcio, A. (2004). A taxonomy of novice user perception of error on the web. *Universal Access in the Information Society, 3*, 202–208.

Lazar, J., Jones, A. & Schneiderman, B. (2006). Workplace user frustration with computers: An exploratory investigation of the causes and severity. *Behaviour & Information Technology, 25 (3)*, 239–251.

Lewis, J. R. (1991). Psychometric evaluation of an after-scenario questionnaire for computer usability studies: The ASQ. *SIHCHI Bulletin, 23(1)*, 78–81.

Lewis, J.R. (1995). IBM computer usability satisfaction questionnaires: Psychometric evaluation and instructions for use. *International Journal of Human-Computer Interaction, 7(1)*, 57–78.

Limann, O. & Pelka, H. (1991). *Televisietechniek*, 2nd edition. Deventer: Kluwer Techniek.

Limesurvey v1.71. An open source survey application, www.limesurvey.org.

Lohr, S.L. (2008). Coverage and sampling. In E.D. de Leeuw, Hox, J.J. & Dillman, D.A. (Eds.), International Handbook of Survey Methodology (pp. 97–112). New York: Lawrence Erlbaum Associates.

Lynn, P. (2008). The problem of nonresponse. In E.D. de Leeuw, Hox, J.J. & Dillman, D.A. (Eds.), *International Handbook of Survey Methodology* (pp. 35–55). New York: Lawrence Erlbaum Associates.

Ma, J. (2007). *Attribution, expectation, and recovery: An integrated model of service failure and recovery*. Doctoral dissertation, Kent State University Graduate School of Management, USA.

Manfreda, K.L. & Vehovar, V. (2008). Internet surveys. In E.D. de Leeuw, Hox, J.J. & Dillman, D.A. (Eds.), *International Handbook of Survey Methodology* (pp. 264– 84). New York: Lawrence Erlbaum Associates.

Mann, H.B. & Whitney, D.R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *Annals of Mathematical Statistics, 18(1)*, 50 – 60.

Martens, J.-B. (2003). *Image technology design – A perceptual approach*. Boston: Kluwer Academic Publishers.

Martin, I. (1991). Expert-novice differences in complaint scripts. *Advances in Consumer Research, 18*, 225–231.

Mendenhall, W. & Sincich, T. (1994). *Statistics for engineering and the sciences*, fourth edition. New Jersey, Prentice-Hall, Inc.

Mitchell, A.A. & Dacin, P.A. (1996). The assessment of alternative measures of consumer expertise. *Journal of Consumer Research, 23(3)*, 219–239.

Moreau, C. P., Lehmann, D.R. & Markman, A.B. (2001). Entrenched knowledge structures and consumer response to new products. *Journal of Marketing Research, 38(1)*, 14–29.

Muller, M., Millen, D.R. & Strohecker, C. (2001). What makes a representative user representative? A participatory poster. *In Proceedings International Conference Extended Abstracts on Human Factors in Computing Systems*, (pp. 101–102), New York: ACM Press.

Mullins, J.W. & Sutherland, D.J. (1998). New product development in rapidly changing markets: An exploratory study. *Journal of Product Innovation Management, 15*, 224–236.

Nielsen, J. (1993). *Usability engineering*. London: Academic Press.

Noldus. Noldus Observer XT 9.0 event logging software.

Norman, D.A. (1983). Some observations on mental models. In D.A. Gentner & A.L. Stevens (Eds.), *Mental models* (pp. 7–14). Hillsdale: Erlbaum.

Norman, D.A. (1998). *The invisible computer.* Cambrige: MIT press.

Norman, D.A. (2002). Home theater: Not ready for prime time. *Computer, 35(6)*, 100–102.

NOS teletekst (2007). Dutch Broadcasting Foundation teletext page retrieved on February 12th 2007.

Novick, D.G. & Ward, K. (2006). Why don't people read the manual? *In Proceedings of the 24th annual ACM international conference on Design of communication*, (pp. 11 – 18), New York: ACM Press.

Oliver, R.L. (1996). *Satisfaction: A behavioral perspective on the consumer*. Boston: Irwin McGraw-Hill.

O'Malley Jr., J.R. (1996). Consumer attributions of product failures to channel members. *Advances in Consumer Research, 23*, 342–346.

Overton, D. (2006, July 16). 'No fault found' returns cost the mobile industry $4.5 billion per year. *WDSGlobal Media Bulletin – White Paper.* Retrieved March 1, 2009, from http://www.wdsglobal.com/news/whitepapers/20060717/MediaBulletinNFF.pdf.

Ozer, M. (1999) . A survey of new product evaluation models. *Journal of Product Innovation Management, 16*, 77–94.

Page, K. & Uncles, M. (1994). Consumer knowledge of the world wide web: Conceptualization and measurement. *Psychology & Marketing, 21(8)*, 573–591.

Park, C.W., & Lessig, V.P. (1981). Familiarity and its impact on consumer decision biases and heuristics. *Journal of Consumer Research, 8*, 223–230.

Park, C.W., Mothersbaugh, D.L. & Feick, L. (1994). Consumer knowledge assessment. *Journal of Consumer Research, 21(1)*, 71–82.

PDMA NPD Glossary (2009). New product development glossary. *Product Development and Management Association*. Retrieved March 2, 2009, from http://www.pdma.org/ npd_glossary.cfm.

Peracchio, L.A. & Tybout, A.M. (1996). The moderating role of prior knowledge in schema-based product evaluation. *The Journal of Consumer Research, 23(3)*, 177–192.

Petkova, V.T. (2003). *An analysis of field feedback in consumer electronics industry*. Doctoral dissertation, Eindhoven University of Technology, The Netherlands.

Pillai, K.G. & Hofacker, C. (2007). Calibration of consumer knowledge of the web. *International Journal of Research in Marketing, 24*, 254–267.

Popovic, V. (2000). Expert and novice user differences and implications for product design and usability. *In Proceedings IEA 2000/HFES 2000 Conference*, (pp. 933–936). San Diego.

Potosnak, K., Hayes, P.J., Rosson, M.B., Schneider, M.L. & Whiteside, J.A. (1986). Classifying users: A hard look at some controversial issues. *In Proceedings of the International Conference on Human Factors in Computing Systems*, (pp. 84–88), Boston: ACM Press.

Pruitt, J. & Grudin, J. (2003). Personas: Practice and theory. *In Proceedings of the 2003 Conference on Designing for User Experiences*, (pp. 1–15), New York: ACM Press.

Prümper, J., Zapf, D., Brodbeck, F.C. & Frese, M. (1992). Some surprising differences between novice and expert errors in computerized work. *Behaviour & Information Technology, 11(6)*, 319–328.

Raju, P.S., Lonial, S.C. & Mangold, W.G. (1995). Differential effects of subjective knowledge, objective knowledge, and usage experience on decision making: An exploratory investigation. *Journal of Consumer Psychology, 4(2)*, 153–180.

Reips, U. (2002a). Standards for internet-based experimenting. *Experimental Psychology, 49(4)*, 243–256.

Reips, U. (2002b). Theory and techniques of conducting web experiments. In B. Batinic, U. Reips & M. Bosnjak (Eds.), *Online social sciences*, (pp. 229–250). Seattle: Hogrefe & Huber Publishers.

Robson, C. (1995). *Real world research: A resource for social scientists and practioner-researchers*. Oxford: Blackwell Publishers.

Rogers, E.M. (2003). *Diffusion of innovations*, 5th edition. New York: Free Press.

Rooden, M.J., & Kanis, H. (2000). Anticipation of usability problems by practitioners. In *Proceedings of the 14th triennial congress of the International Ergonomics Association and 44th annual meeting of the Human factors and Ergonomics Society*, (pp. 6-941-6-944). Santa Monica CA (USA): Human Factors and Ergonomics Society.

Rooijmans, J., Aerts, H. & Genuchten, M. (1996). Software quality in consumer electronics products. *IEEE Software, 13*, 55–64.

Roth, E.M., Patterson, E.S. & Mumaw, R.J. (2002). Cognitive engineering: Issues in user-centered systems design. In J.J. Marciniak (Ed.), *Encyclopedia of software engineering*, second edition (pp. 163–179). New York: John Wiley & Sons Inc.

Rozinat, A. & Van der Aalst, W.M.P. (2008). Conformance checking of processes based on monitoring real behavior. *Information Systems, 33(1)*, 64–95.

Russell, D. (1982). The causal dimension scale: A measure of how individuals perceive causes. *Journal of Personality and Social Psychology, 42(6)*, 1137–1145.

Russell, D. (1987). Measuring causal attributions for success and failure: A comparison of methodologies for assessing causal dimensions. *Journal of Personality and Social Psychology 52(6)*, 1248–1257.

Rust, R.T., Thompson, D.V. & Hamilton, R.W. (2006). Defeating Feature Fatigue. *Harvard Business Review, 84 (2)*, 98–107.

Saaksjarvi, M. (2003). Consumer adoption of technological innovations. *European Journal of Innovation Management, 6(2)*, 90–100.

Schwarz, N., Knäuper, B., Oyserman, D. & Stich, C. (2008). The psychology of asking questions. In E.D. de Leeuw, Hox, J.J. & Dillman, D.A. (Eds.), *International Handbook of Survey Methodology* (pp. 264–284). New York: Lawrence Erlbaum Associates.

Senternovem (2005). Factsheet 'managing soft reliability' project. IOP integral product creation and realization. Den Haag: Senternovem. Retrieved April 30, 2009, from http://www.senternovem.nl/mmfiles/Factsheet%20managing%20soft%20reliability-EN_tcm24-277605.pdf.

Senternovem (2008). Factsheet 'datafusion' project. IOP integral product creation and realization. Den Haag: Senternovem. Retrieved April 30, 2009 from, http://www.senternovem.nl/mmfiles/Factsheet%20data%20fusion-EN_tcm24-277589.pdf.

Shackel, B. (1984). The concept of usability. In J. Bennet, J., D. Case, J. Sandelin, &M. Smith, (Eds.), *Visual Display Terminals* (pp. 45–87). New Jersey: Prentice-Hall.

Sharma, A. & Silver, S. (2009, January 26). Blackberry Storm is off to a bit of a bumpy start. *The Wall Street Journal*. Retrieved March 1, 2009, from http://online.wsj.com/article/SB123292905716613927.html.

Shih, C-F. & Venkatesh, A. (2004). Beyond adoption: Development and application of a use-diffusion model. *Journal of Marketing, 68*, 59–72.

Siegel, S. (1957). Nonparametric statistics. *The American Statistician, 11(3)*, 13–19.

Siewiorek, D.P., Chillarge, R. & Kalbarczyk, Z.T. (2004). Reflections on industry trends and experimental research in dependability. *IEEE Transactions on Dependable and Secure Computing, 2(1)*, 109–127.

Silvera, D. H. & Laufer, D. (2005). Recent developments in attribution research and their implications for consumer judgments and behavior. In F.R. Kardes, P.M. Herr& J. Nantel (Eds.). *Applying social cognition to consumer-focused strategy* (pp. 53–80). Mahwah: Lawrence Erlbaum Associates Inc.

Smith, A.K., Bolton, R.N. & Wagner, J. (1999). A model of customer satisfaction with service encounters involving failure and recovery. *Journal of Marketing Research, 36 (3)*, 356–372.

Smith, B., Caputi, P., Crittenden, N., Jayasuriya, R. & Rawstorne, P. (1999). A review of the construct of computer experience. *Computers in Human Behavior, 15*, 227–242.

Smith, E.E., Nolen-Hoeksma, S.N., Fredrickson, B.L. & Loftus, G.R. (2003). *Atkinson & Hilgard's Introduction to Psychology*, 14th edition. Belmont: Wadsworth/Thomson Learning.

Söderlund, M. (2002). Customer familiarity and its effects on satisfaction and behavioural intentions. *Psychology & Marketing, 19(10)*, 861–880.

Somasundaram, T.N. (1993). Consumer reaction to product failure: Impact of product involvement and knowledge. *Advances in Consumer Research, 20*, 215–218.

Song, M. & Van der Aalst, W.M.P. (2007). Supporting process mining by showing events at a glance. *In K. Chari, A. Kumar (Eds.), Seventeenth Annual Workshop on Information Technologies and Systems (WITS'07)* (pp 139–145), Montreal, Canada, December 8-9, 2007.

Staggers, N. & Norcio, A.F. (1993). Mental models: Concepts for human-computer interaction research. *International Journal of Man-Machine Studies, 38*, 587–605.

Stangor, C. (1998). *Research methods for the behavioral sciences*. Boston: Houghton Mifflin Company.

Steger, T., Sprague, B. & Douthit, D. (2007). Big trouble with no trouble found: How consumer electronics firms confront the high cost of customer returns. *Accenture Communications & High Tech.*

Stroucken, L., Seeverens H., Beenker, F. & Watts, D. (2005). *Trader, television related architecture and design to enhance reliability*. Summary project plan, public version, (ESI document No. 2005-11117, version 1): Embedded Systems Institute Eindhoven. Retrieved February 27, 2009, from http://www.esi.nl/projects/trader/publications/summary_trader_projectplan_public.pdf

Sujan, M. (1985). Consumer knowledge: Effects on evaluation strategies mediating consumer judgments. *Journal of Consumer Research, 12(1)*, 31–46.

Taylor, S. & Todd, P.A. (1995). Understanding information technology usage: A test of competing models. *Information System Research, 6(2)*, 144–176.

Tekinerdogan, B., Sözer, H. & Aksit, M. (2008). Software architecture reliability analysis using failure scenarios. *Journal of Systems and Software, 81(4)*, 558–575.

Thatcher, A & Greyling, M. (1998). Mental models of the Internet. *International Journal of Industrial Ergonomics 22*, 299–305.

Thompson, D.V., Hamilton, R.W. & Rust, R.T. (2005). Feature fatigue: When product capabilities become too much of a good thing. *Journal of Marketing Research, 42(4)*, 431–442.

Tronvoll, B. (2007). Customer complaint behavior from the perspective of the service-dominant logic of marketing. *Managing Service Quality, 17(6)*, 601–620.

Uther, M. & Haley, H. (2006). Back vs. stack: Training the correct mental model affects web browsing. *Behaviour and Information Technology, 27(3)*, 211–218.

Van den Ende, N., De Hesselle, H. & Meesters, L. (2007). Towards content-aware coding: User study. *In P. Cesar, K. Chorianopoulos, & J.F. Jensen (eds.) EuroITV 2007. LNCS, vol. 4471*, (pp. 185–194). Heidelberg: Springer.

Van der Aalst, W.M.P., Van Dongen, B.F., Günther, C.W., Mans, R.S., Alvas de Medeiros, A.K., Rozinat, A., Rubin, V., Song, M., Verbeek, H.M.W. & Weijters, A.J.M.M. (2007). ProM 4.0: Comprehensive support for *real* process analysis. In J. Kleijn & A. Yakovlev. *Petri Nets and Other Models of Concurrency* (Proceedings 28th International Conference on Applications and Theory of Petri Nets and Other Models of Concurrency, ICATPN 2007, Siedcle, Poland, June 25-29, 2007). *Lecture Notes in Computer Science, 4546* (484–494). Berlin: Springer.

Van der Veer, G.C. & Del Carmen Puerta Melguizo, M. (2002). Mental models. In J.A. Jacko & A. Sears (Eds.), *The human-computer interaction handbook: Fundamentals, evolving technologies and emerging applications* (pp. 52–80). Hillsdale: Lawrence Erlbaum Associates Inc.

Van Dijk, W. (2008, August 12). T-Mobile reageert op klachten decking iPhone. Retrieved March 19, 2009, from http://www.nu.nl/internet/1697519/t-mobile-reageert-op-klachten-dekking-iphone.html.

Varma, S. (n.d.). *Preliminary item statistics using point-biserial correlation and P-values*. Education Data Systems Inc, California, Morgan Hill. Retrieved August 18, 2009, from http://www.eddata.com/resources/publications/EDS_Point_Biserial.pdf.

Venkatesh, V, Morris, M.G., Davis, G.B. & Davis, F.D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly, 27(3)*, 425–478.

Verbeek, P.-P. & Slob., A. (2006). Analyzing the relations between technology and user behavior: Towards a conceptual framework. In P.-P. Verbeek & A. Slob (Eds.), *User behavior and technology development: Shaping relations between consumers and technologies* (pp. 385–399). Dordrecht: Springer.

Von Hippel, E. (1986). Lead users: A source of novel product concepts. *Management Science, 32(7)*, 791–805.

Vredenburg, K., Isensee, S. & Righi, C. (2002). *User-centered design: An integrated approach*. New Jersey: Prentice Hall.

Weijters, A.J.M.M. & Van der Aalst, W.M.P. (2003). Rediscovering workflow models from event-based data using Little thumb. *Integral Computer-Aided Engineering, 10(2)*, 151–162

Weiner, B. (1985). An attributional theory of achievement, motivation and emotion. *Psychological Review 92*, 548–573.

Weiner, B. (1986). *An attributional theory of motivation and emotion*. New York: Springer-Verlag.

Weiner, B. (2000). Attributional thoughts about consumer behavior. *Journal of Consumer Research, 27*, 382–387.

Wever, R., Van Kuijk, J & Boks, C. (2008). User-centred design for sustainable behaviour. *International Journal of Sustainable Engineering, 1(1)*, 9–20.

Yasper (2009).*Yet Another Smart Process EditoR*. Software for modelling workflow processes using extended Petri Nets. Eindhoven University of Technology and Deloitte. Downloaded from http://www.yasper.org.

Yin, R.K. (1994). *Case study research – Design and methods*, second edition. London: Sage Publications Ltd.

YouTube. Video sharing website, www.youtube.com.

Zaichkowsky, J. L. (1985). Measuring the involvement construct. *The Journal of Consumer Research, 12(3)*, 34 –352.

Zhang, X. & Chignell, M. (2001). Assessment of the effects of user characteristics on mental models of information retrieval systems. *Journal of the American Society for Information Science and Technology, 52(6)*, 445–459.

Ziefle, M. (2002). The influence of user expertise and phone complexity on performance, ease of use and learnability of different mobile phones. *Behaviour & Information Technology, 21(5)*, 303–311.

Ziefle, M. & Bay, S. (2004). Mental models of cellular phone menu. Comparing older and younger novice users. In *Lecture notes in computer science: Mobile human-computer interaction – MobileHCI 2004* (pp. 25–37). Berlin: Springer.

Appendices Chapter 3

Appendix 3.1 Measurement of teletext usage experience (in Dutch)[28]

Welke teletekstfuncties gebruikt u en hoe vaak gebruikt u die gemiddeld?

Teletekstfunctie	Gebruiksfrequentie					
	Dagelijks meerdere keren	Dagelijks	Wekelijks	Een enkele keer	Nooit	Weet niet
Nieuws	☐	☐	☐	☐	☐	☐
Televisiegids	☐	☐	☐	☐	☐	☐
Radiogids	☐	☐	☐	☐	☐	☐
Omroepinformatie (zoals Tros, BNN etc.)	☐	☐	☐	☐	☐	☐
Shownieuws / filmnieuws / muziek / uitgaan	☐	☐	☐	☐	☐	☐
Financiën	☐	☐	☐	☐	☐	☐
Sport	☐	☐	☐	☐	☐	☐
Weer	☐	☐	☐	☐	☐	☐
Reizen	☐	☐	☐	☐	☐	☐
Ondertiteling	☐	☐	☐	☐	☐	☐
Overig, namelijk _____	☐	☐	☐	☐	☐	☐
_____	☐	☐	☐	☐	☐	☐
_____	☐	☐	☐	☐	☐	☐

[28] An (not validated) English translation of the questionnaires in this Appendix can be obtained from the author.

Op welke televisiekanalen gebruikt u teletekst en hoe vaak?

Televisiekanaal	**Gebruiksfrequentie teletekst**					
	Dagelijks meerdere keren	Dagelijks	Wekelijks	Een enkele keer	Nooit	Weet niet
Nederland 1, 2, 3	☐	☐	☐	☐	☐	☐
RTL 4, 5, 7	☐	☐	☐	☐	☐	☐
SBS 6 / Net 5 / Veronica	☐	☐	☐	☐	☐	☐
TMF / MTV / The Box	☐	☐	☐	☐	☐	☐
Talpa	☐	☐	☐	☐	☐	☐
Lokale zender	☐	☐	☐	☐	☐	☐
Belgische zenders	☐	☐	☐	☐	☐	☐
Duitse zenders	☐	☐	☐	☐	☐	☐
Engelse zenders	☐	☐	☐	☐	☐	☐
Overig, namelijk _____	☐	☐	☐	☐	☐	☐
_____	☐	☐	☐	☐	☐	☐

Appendix 3.2 Task list teletext experiment

Takenlijst TRADER consumententest teletekst gebruiksvriendelijkheid

Hieronder staan drie scenario's beschreven waarin informatie van teletekst nodig is. Wij willen je vragen om in de onderstaande volgorde de benodigde informatie op te zoeken in teletekst op Nederland 1, 2 of 3. Probeer hierbij net te doen alsof je je TV thuis ook gebruikt.

Wij willen je vragen om alles wat je doet en denkt hardop te zeggen zodat wij kunnen volgen wat je doet en denkt. Mocht je er in het uiterste geval niet uit komen kun je de observatoren vragen om hulp.

Om de informatie met betrekking tot het experiment op te kunnen slaan willen wij je erop wijzen dat je de TV net als thuis kunt gebruiken behalve dat je de TV niet uit mag zetten en niet aan de kabels van de TV mag komen!

Probeer om de volgende informatie zo goed en volledig mogelijk op te zoeken in teletekst:

1. PSV scenario

Je bent een grote fan van het voetbalteam van PSV Eindhoven. Je bent net terug van vakantie en je wilt graag weten:

- Wanneer PSV de volgende wedstrijd moet spelen
- Op welke positie PSV op de ranglijst staat

2. TV-gids scenario

Je hebt thuis geen TV-gids en je wilt graag weten of er vanavond leuke films zijn te zien op de televisiekanalen RTL5 en SBS6. Zoek op welke films er vanavond te zien zijn op RTL5 en SBS6.

3. Schiphol scenario

Je moet je partner afhalen op Schiphol en je weet niet hoe laat zijn/haar vliegtuig aankomt. Je weet dat hij/zij aankomt met een vlucht uit Helsinki met vluchtnummer KL 1168. Zoek op hoe laat het vliegtuig aankomt.

Bedankt voor je tijd!

Appendices Chapter 4

Appendix 4.1 Questionnaire items (Dutch translation)[29]

Construct	Items	Code	Response scale
Subjective expertise televisions	1. Ik weet vrij veel van televisies 2. Ik heb niet het gevoel veel te weten van televisies 3. In mijn vriendenkring ben ik een van de 'experts' op het gebied van televisies 4. Vergeleken met de meeste andere mensen weet ik weinig van televisies 5. Wat televisies betreft weet ik niet erg veel	SubExTel_1 SubExTel_2 SubExTel_3 SubExTel_4 SubExTel_5	5 point likert scale: • Mee eens • Enigszins mee eens • Niet mee eens, niet mee oneens • Enigszins mee oneens • Mee oneens
Television usage (objective)	Hoeveel uur maakt u, over het algemeen, gemiddeld per week gebruik van uw televisie?	UseTel_1	5 point scale: • Meer dan 20 uur per week • 15 – 20 uur per week • 10 – 15 uur per week • 5 – 10 uur per week • Minder dan 5 uur per week
Familiarity: information search for television purchase	1. Als ik overweeg een televisie te kopen raadpleeg ik meerdere informatiebronnen 2. Als ik een televisie wil kopen, doe ik vergeleken met de meeste andere mensen niet veel moeite om informatie op te zoeken 3. Als ik een televisie koop, zoek ik goed advies voordat ik een beslissing neem	SearchPTel_1 SearchPTel_2 SearchPTel_3	5 point likert scale
Familiarity: information search during television usage	1. Ik maak weinig gebruik van de handleiding van mijn televisie om informatie op te zoeken over mijn televisie 2. Ik zoek vaak naar informatie op het Internet over het gebruik van mijn televisie 3. Als ik iets wil weten over mijn televisie vraag ik vaak om hulp aan vrienden, collega's of de helpdesk	SearchUseTel_1 SearchUseTel_2 SearchUseTel_3	5 point likert scale
Familiarity: television usage (subjective)	1. Vergeleken met de meeste andere mensen gebruik ik mijn televisie vaak 2. In mijn dagelijks leven kan ik niet zonder mijn televisie 3. Vergeleken met de meeste andere mensen maak ik weinig gebruik van mijn televisie	UseTel_2 UseTel_3 UseTel_4	5 point likert scale

[29] An (not validated) English translation of the questionnaires in this Appendix can be obtained from the author.

Construct	Items	Code	Response scale
Familiarity: general information search on televisions	1. Ik zoek regelmatig naar informatie over nieuwe ontwikkelingen op het gebied van televisies, ook al heb ik geen nieuwe televisie nodig 2. Ik lees graag over televisies in tijdschriften of op het Internet, omdat het me interesseert. 3. Ik praat regelmatig met mijn vrienden of collega's over nieuwe ontwikkelingen op het gebied van televisies	SearchInfoTel_1 SearchInfoTel_2 SearchinfoTel_3	5 point likert scale
Subjective expertise computers	Same as for subjective expertise televisions (replace 'televisies' with 'computers')	SubExCom_1 – SubExCom_5	5 point likert scale
Computer usage (objective)	Same as for television usage (replace 'televisie' with 'computer')	UseCom_1	Same as for television usage (objective)
Familiarity: information search for computer purchase	Same as for information search for television purchase (replace 'televisies' with 'computers')	SearchPCom_1 – SearchPCom_3	5 point likert scale
Familiarity: information search during computer usage	Same as for information search during television purchase (replace 'televisies' with 'computers')	SearchUSeCom_1 – SearchUseCom_3	5 point likert scale
Familiarity: computer usage (subjective)	Same as for television usage (subjective) (replace 'televisies' with 'computers')	UseCom_2 – UseCom_4	5 point likert scale
Familiarity: general information search on computers	Same as for general information search on televisions (replace 'televisies' with 'computers')	SearchInfoCom_1 – SearchInfoCom_3	5 point likert scale
Intention-to-use multimedia LCD televisions introduction	Introduction text: De volgende stellingen hebben betrekking op televisies met multimedia toepassingen. Naast het kijken van televisieprogramma's en teletekst hebben deze innovatieve televisies ook andere functionaliteiten zoals een digitale media reader en multimedia-aansluitingen. Met een digitale media reader kunt u bijvoorbeeld digitale foto's bekijken op de televisie. Met multimedia-aansluitingen kunt u, naast het aansluiten van de televisie op bijvoorbeeld een videorecorder of DVD speller, de televisie (draadloos) aansluiten op uw computer om vervolgens foto's, muziek of films van uw computer op uw televisie af te spleen.	X	X

Construct	Items	Code	Response scale
Intention-to-use: performance expectancy	1. Ik zou een televisie met multimedia toepassingen nuttig vinden in mijn leven 2. Ik verwacht dat ik door het gebruik van een televisie met multimedia toepassingen in staat ben dingen sneller te doen 3. Het gebruik van een televisie met multimediatoepassingen zou het voor mij mogelijk maken om meer dingen te doen in dezelfde tijd	PerfExpMMTV_1 PerfExpMMTV_2 PerfExpMMTV_3	5 point likert scale
Intention-to-use: effort expectancy	1. Ik verwacht dat mijn interactie met een televisie met multimedia toepassingen duidelijk en begrijpelijk zou zijn 2. Het zou gemakkelijk voor mij zijn om het gebruik van een televisie met multimediatoepassingen onder de knie te krijgen 3. Ik zou een televisie met multimedia toepassingen gemakkelijk te gebruiken vinden 4. Het zou gemakkelijk zijn voor mij om een televisie met multimedia toepassingen te leren bedienen	EffExpMMTV_1 EffExpMMTV_2 EffExpMMTV_3 EffExpMMTV_4	5 point likert scale
Intention-to-use: facilitating conditions	1. Ik verwacht dat ik de benodigde kennis heb om een televisie met multimedia toepassingen te gebruiken 2. Ik verwacht date en televisie met multimedia toepassingen niet te gebruiken is in combinatie met andere apparaten die ik gebruik 3. Ik ken iemand die mij zou helpen als ik problemen met een televisie met multimediatoepassingen zou hebben	FacConMMTV_1 FacConMMTV_2 FacConMMTV_3	5 point likert scale

Appendix 4.2 Final solution for the basic factor analysis

Table A4.1 MSA and Bartlett's test for the final factor solution for the basic factor analysis

Kaiser-Meyer-Olkin Measure of Sampling Adequacy		0.877
Bartlett's Test of Sphericity	Approx. Chi-Square	4962.537
	df	300
	Sig.	.000

Table A4.2 Pattern matrix of the basic factor analysis

	Components					
	1	**2**	**3**	**4**	**5**	**6**
SubExTel_1				0.761		
SubExTel_2				0.819		
SubExTel_3				0.650		
SubExTel_4				0.878		
SubExTel_5				0.907		
SearchPTel_1		0.941				
SearchPTel_2		0.884				
SearchPTel_3		0.883				
UseTel_3					0.854	
UseTel_4					0.807	
SearchInfoTel_1	0.873					
SearchInfoTel_2	0.932					
SearchInfoTel_3	0.796					
SubExCom_1			-0.457			
SubExCom_3	0.410					
SubExCom_4			-0.585			
SearchPCom_1						-0.856
SearchPCom_2						-0.810
SearchPCom_3						-0.897
UseCom_2			-0.890			
UseCom_3			-0.883			
UseCom_4			-0.894			
SearchInfoCom_1	-0.738					
SearchInfoCom_2	-0.767					
SearchInfoCom_3	-0.708					

Extraction method: Principal components analysis
Rotation method: Oblimin with Kaiser Normalization
Rotation converged in nine iterations

Appendix 4.3 Factor analysis of intention-to-use

Table A4.3 MSA and Bartlett's test for the final factor solution of intention-to-use

Kaiser-Meyer-Olkin Measure of Sampling Adequacy		0.853
Bartlett's Test of Sphericity	Approx. Chi-Square	1014.328
	df	21
	Sig.	.000

Table A4.4 Pattern matrix of the final factor analysis of intention-to-use

	Components	
	1	2
PerfExpMMTV_1		0.643
PerfExpMMTV_2		0.964
PerfExpMMTV_3		0.940
EffExpMMTV_1	0.525	
EffExpMMTV_2	0.866	
EffExpMMTV_3	0.918	
EffExpMMTV_4	0.943	

Extraction method: Principal components analysis
Rotation method: Oblimin with Kaiser Normalization
Rotation converged in six iterations

Appendices Chapter 5

Appendix 5.1 Introduction to experiment (in Dutch)[30]

De test die u zodadelijk gaat uitvoeren heeft betrekking op de gebruiksvriendelijkheid van de LCD televisie die voor u staat. Stelt u zich voor dat u deze televisie enkele dagen geleden heeft gekocht. Voor deze televisie heeft u €3250,- betaald. Inmiddels zijn alle kabels en stekkers op de televisie aangesloten en zijn de TV zenders met behulp van de automatische installatie geïnstalleerd.

In de winkel heeft u gehoord van de verkoper dat deze televisie meer functies heeft dan uw oude televisie. U wil deze functies graag uitproberen om te zien wat u ermee kunt. Op de volgende pagina's staan drie verschillende taken die te maken hebben met de innovatieve functies van deze televisie.

Wij willen u voor deze test vragen elke taak uit te voeren. Bij het uitvoeren van deze taken vragen wij u op de volgende dingen te letten:

- Op elke pagina staat één taak beschreven. Graag eerst de taak uitvoeren en daarna pas naar de volgende pagina gaan.

- Tijdens het uitvoeren van de taak willen wij u vragen **hardop na te denken** over elke stap die u neemt om de taak uit te voeren. Dit zorgt ervoor dat wij uw denkproces tijdens deze test kunnen volgen.

- Om de data van de test op te kunnen slaan willen wij u vragen **niet** aan de kabels van de televisie te komen.

Tot slot willen wij u erop wijzen dat het tijdens deze test niet gaat om iets goed of fout doen. Wij zijn geïnteresseerd in uw ervaring met deze televisie om producten in de toekomst te verbeteren.

[30] An (not validated) English translation of the questionnaires in this Appendix can be obtained from the author.

Appendix 5.2 Task list (in Dutch)

<u>Gebruik van dual screen</u> (dual screen task)

De "dual screen" functie van deze televisie biedt u de mogelijkheid om tegelijkertijd twee TV kanalen op het scherm van de televisie weer te geven.

Voor deze taak vragen wij u tegelijkertijd de zenders "Nederland 2" en "Net 5" op het televisiescherm te laten zien met behulp van de "dual screen" functie.

Vul na het uitvoeren van deze taak de bijbehorende vragenlijst in.

<u>Een TV kanaal herprogrammeren</u> (switch channel task)

Met behulp van de automatische installatie zijn de TV zenders op de televisie ingesteld. Nederland 1 staat op kanaal 1, Nederland 2 op kanaal 2, RTL4 op kanaal 4 enzovoorts. Op uw oude televisie had u RTL7 op kanaal 7 staan, maar de automatische installatie heeft deze zender op deze televisie op een ander kanaal geprogrammeerd. U wilt graag de zenders op deze televisie op dezelfde volgorde programmeren als op uw oude televisie.

Voor deze taak vragen wij u om RTL7 op kanaal 7 te programmeren.

Vul na het uitvoeren van deze taak de bijbehorende vragenlijst in.

<u>Digitale foto's bekijken</u> (digital picture task)

De verkoper in de winkel heeft u verteld dat u op deze televisie digitale foto's kunt bekijken die op een USB stick of geheugenkaart staan. Omdat u vaak digitale foto's maakt, wilt u deze functie graag uitproberen.

Voor deze taak vragen wij u om de digitale foto's die op de USB stick staan, te bekijken op de televisie.

Vul na het uitvoeren van deze taak de bijbehorende vragenlijst in.

Appendix 5.3 Task exit questionnaire (in Dutch)

Onderstaande vragen hebben betrekking op het gebruik van de (dual screen / TV kanaal / digitale foto) functie.

Omcirkel bij elke vraag het antwoord dat het beste uw mening weergeeft.

1. Over het algemeen ben ik tevreden over het gemak waarmee ik deze taak kon uitvoeren.

Volledig mee oneens	1	2	3	4	5	6	7	Volledig mee eens

2. Over het algemeen ben ik tevreden over de snelheid waarmee ik deze taak kon uitvoeren.

Volledig mee oneens	1	2	3	4	5	6	7	Volledig mee eens

3. Over het algemeen ben ik tevreden over de beschikbare ondersteunende informatie (handleiding, informatie op het televisiescherm) voor het uitvoeren van deze taak.

Volledig mee oneens	1	2	3	4	5	6	7	Volledig mee eens

4. Heeft uw eigen televisie thuis een (vergelijkbare) functie die u bij deze taak heeft gebruikt?

☐ Ja

☐ Nee

☐ Weet niet

5. Heeft u ervaring met het gebruik van deze functie (of een vergelijkbare functie) op een televisie?

☐ Ja

☐ Nee

☐ Weet niet

Appendix 5.4 Participants characteristics based on differentiation on subjective expertise of computers

Table A5.1 Overview of participant characteristics based on differentiation on subjective expertise of computers

	High SubExCom (n = 15)			Low SubExCom (n = 14)		
	mean	**S.D.**	**range**	**mean**	**S.D.**	**range**
Age	35.67	12.51	22 – 59	52.00	12.04	31 – 67
Intention-to-use	4.21	0.83	2.29 – 5.00	3.42	1.07	1.00 – 4.86
SubExTel	3.95	1.07	1.80 – 5.00	1.83	0.48	1.00 – 2.60
SubFamTel	2.99	1.14	0.75 – 4.38	2.83	0.79	1.13 – 3.88
ObjFamTel	3.27	1.39	1.00 – 5.00	4.00	0.96	2.00 – 5.00
SubExCom	4.32	0.76	2.80 – 5.00	1.89	0.49	1.00 – 2.60
SubFamCom	3.76	1.18	1.00 – 5.00	2.58	1.06	0.33 – 3.56
ObjFamCom	4.33	1.40	1.00 – 5.00	3.64	1.60	0.00 – 5.00

High SubExCom group: 13 males and 2 females
Low SubExCom group: 6 males and 8 females

Results of separate pair-wise Mann Whitney U tests showed significant differences between the knowledge groups on subjective expertise of computers ($p < 0.001$), subjective expertise of televisions ($p < 0.001$), subjective familiarity with computers ($p < 0.01$), intention-to-use ($p < 0.05$), age ($p < 0.01$) and gender ($p < 0.05$).

Appendix 5.5 MANOVA results for evaluation of the effect of subjective expertise

Table A5.2 Multivariate tests with subjective expertise of televisions

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial eta squared	Noncent. Parameter	Observed power
Intercept	Pillai's Trace	.977	934.650	3	66	.000	.977	2803.949	1.000
	Wilks' Lambda	.023	934.650	3	66	.000	.977	2803.949	1.000
	Hotelling's Trace	42.484	934.650	3	66	.000	.977	2803.949	1.000
	Roy's Largest Root	42.484	934.650	3	66	.000	.977	2803.949	1.000
Task complexity	Pillai's Trace	.371	5.080	6	134	.000	.185	30.479	.992
	Wilks' Lambda	.630	5.717	6	132	.000	.206	34.300	.997
	Hotelling's Trace	.586	6.350	6	130	.000	.227	38.102	.999
	Roy's Largest Root	.584	13.052	6	67	.000	.369	39.156	1.000
SubExTel group	Pillai's Trace	.261	7.773	3	66	.000	.261	23.318	.985
	Wilks' Lambda	.739	7.773	3	66	.000	.261	23.318	.985
	Hotelling's Trace	.353	7.773	3	66	.000	.261	23.318	.985
	Roy's Largest Root	.353	7.773	3	66	.000	.261	23.318	.985
Task complexity * SubExTel group	Pillai's Trace	.116	1.380	6	134	.227	.058	8.277	.525
	Wilks' Lambda	.886	1.379	6	132	.228	.059	8.273	.524
	Hotelling's Trace	.127	1.377	6	130	.228	.060	8.265	.523
	Roy's Largest Root	.107	2.399	6	67	.076	.097	7.197	.575

Table A5.3 Test of between-subjects effects for MANOVA with subjective expertise of TVs

Source	Dependent variable	Type III sum of squares	df	Mean square	F	Sig.	Partial eta squared	Observed power[b]
Corrected model	Time	719.346[a]	5	143.869	5.010	.001	.269	.976
	Levelup	23.292[b]	5	4.658	5.628	.000	.293	.988
	Steps	2.854[c]	5	.571	8.277	.000	.378	.999
Intercept	Time	18160.867	1	18160.867	632.398	.000	.903	1.000
	Levelup	111.078	1	111.078	134.195	.000	.664	1.000
	Steps	75.549	1	75.549	1095.536	.000	.942	1.000
Task complexity	Time	206.584	2	103.292	3.597	.033	.096	.648
	Levelup	16.210	2	8.105	9.792	.000	.224	.979
	Steps	2.340	2	1.170	16.969	.000	.333	1.000
SubExTel group	Time	525.087	1	525.087	18.285	.000	.212	.988
	Levelup	2.735	1	2.735	3.304	.074	.046	.433
	Steps	.131	1	.131	1.903	.172	.027	.275
Task complexity * SubExTel group	Time	69.190	2	34.595	1.205	.306	.034	.255
	Levelup	5.995	2	2.997	3.621	.032	.096	.651
	Steps	.448	2	.224	3.252	.045	.087	.601
Error	Time	1952.788	68	28.717				
	Levelup	56.286	68	.828				
	Steps	4.689	68	.069				
Total	Time	20503.000	74					
	Levelup	177.000	74					
	Steps	81.145	74					
Corrected total	Time	2672.134	73					
	Levelup	79.578	73					
	Steps	7.543	73					

a R squared = 0.269 (Adjusted R squared = 0.215)

b Computed using alpha = 0.05

c R squared = 0.293 (Adjusted R squared = 0.241)

d R squared = 0.378 (Adjusted R squared = 0.333)

Table A5.4 Multivariate tests with subjective expertise of computers

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial eta squared	Noncent. Parameter	Observed power
Intercept	Pillai's Trace	.975	869.077	3	66	.000	.975	2607.323	1.000
	Wilks' Lambda	.025	869.077	3	66	.000	.975	2607.323	1.000
	Hotelling's Trace	39.504	869.077	3	66	.000	.975	2607.323	1.000
	Roy's Largest Root	39.504	869.077	3	66	.000	.975	2607.323	1.000
Task complexity	Pillai's Trace	.346	4.678	6	134	.000	.173	28.066	.987
	Wilks' Lambda	.655	5.179	6	132	.000	.191	31.071	.993
	Hotelling's Trace	.524	5.674	6	130	.000	.208	34.046	.997
	Roy's Largest Root	.519	11.594	6	67	.000	.342	34.782	.999
SubExCom group	Pillai's Trace	.304	9.591	3	66	.000	.304	28.774	.996
	Wilks' Lambda	.655	9.591	3	66	.000	.304	28.774	.996
	Hotelling's Trace	.436	9.591	3	66	.000	.304	28.774	.996
	Roy's Largest Root	.436	9.591	3	66	.000	.304	28.774	.996
Task complexity * SubExCom group	Pillai's Trace	.073	.846	6	134	.537	.036	5.075	.326
	Wilks' Lambda	.928	.844	6	132	.539	.037	5.061	.325
	Hotelling's Trace	.078	.841	6	130	.541	.037	5.045	.323
	Roy's Largest Root	.070	1.565	6	67	.206	.065	4.694	.394

Table A5.5 Test of between-subjects effects for MANOVA with subjective expertise of computers

Source	Dependent variable	Type III sum of squares	df	Mean square	F	Sig.	Partial eta squared	Observed power[b]
Corrected model	Time	753.829[a]	5	150.766	5.344	.000	.282	.983
	Levelup	20.990[c]	5	4.198	4.872	.001	.264	.973
	Steps	2.608[d]	5	.522	7.188	.000	.346	.998
Intercept	Time	17918.613	1	17918.613	635.178	.000	.903	1.000
	Levelup	110.272	1	110.272	127.988	.000	.653	1.000
	Steps	72.034	1	72.034	992.590	.000	.936	1.000
Task complexity	Time	239.915	2	119.958	4.252	.018	.111	.725
	Levelup	17.131	2	8.566	9.942	.000	.226	.981
	Steps	2.293	2	1.146	15.795	.000	.317	.999
SubExCom group	Time	571.272	1	571.272	20.250	.000	.229	.993
	Levelup	2.616	1	2.616	3.036	.086	.043	.404
	Steps	.104	1	.104	1.431	.236	.021	.218
Task complexity * SubExCom group	Time	30.795	2	15.398	.546	.582	.016	.137
	Levelup	3.906	2	1.953	2.267	.111	.063	.446
	Steps	.226	2	.113	1.558	.218	.044	.320
Error	Time	1918.305	68	28.210				
	Levelup	58.588	68	.862				
	Steps	4.935	68	.073				
Total	Time	20503.000	74					
	Levelup	177.000	74					
	Steps	81.145	74					
Corrected total	Time	2672.134	73					
	Levelup	79.578	73					
	Steps	7.543	73					

a R squared = 0.282 (Adjusted R squared = 0.229)

b Computed using alpha = 0.05

c R squared = 0.264 (Adjusted R squared = 0.210)

d R squared = 0.346 (Adjusted R squared = 0.298)

Appendix 5.6 Desired usage models for the tasks

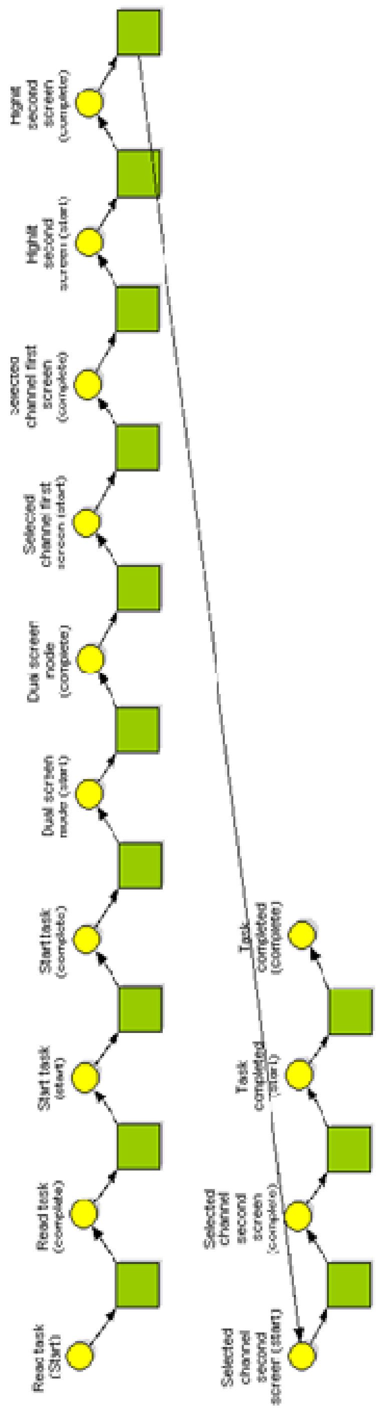


Figure A5.1 Designed usage model of dual screen task

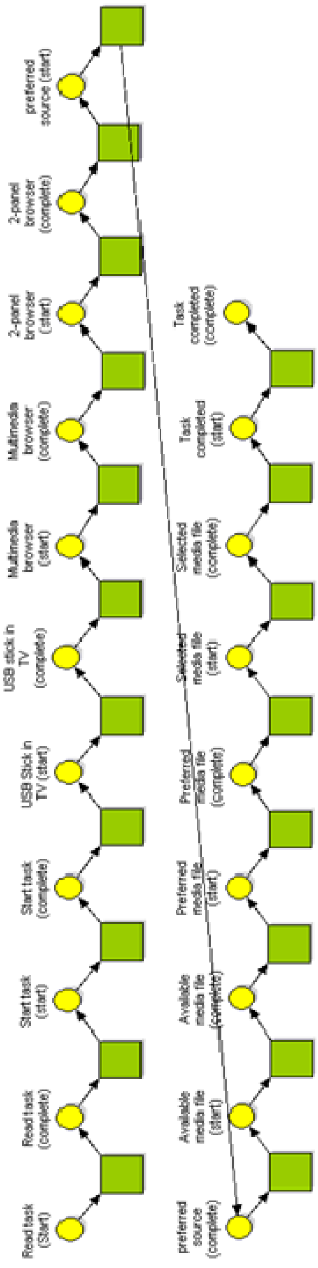


Figure A5.2 Designed usage model of digital picture task

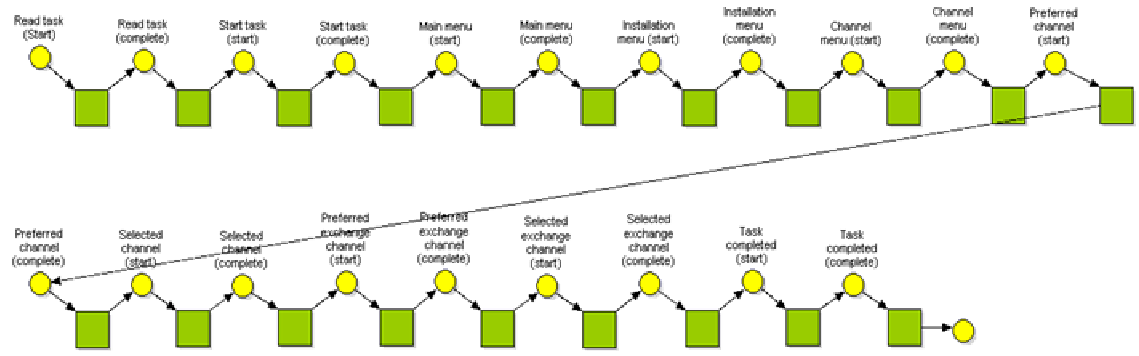


Figure A5.3 Designed usage model of switch channel task

Appendix 5.7 Mined heuristic models of the tasks

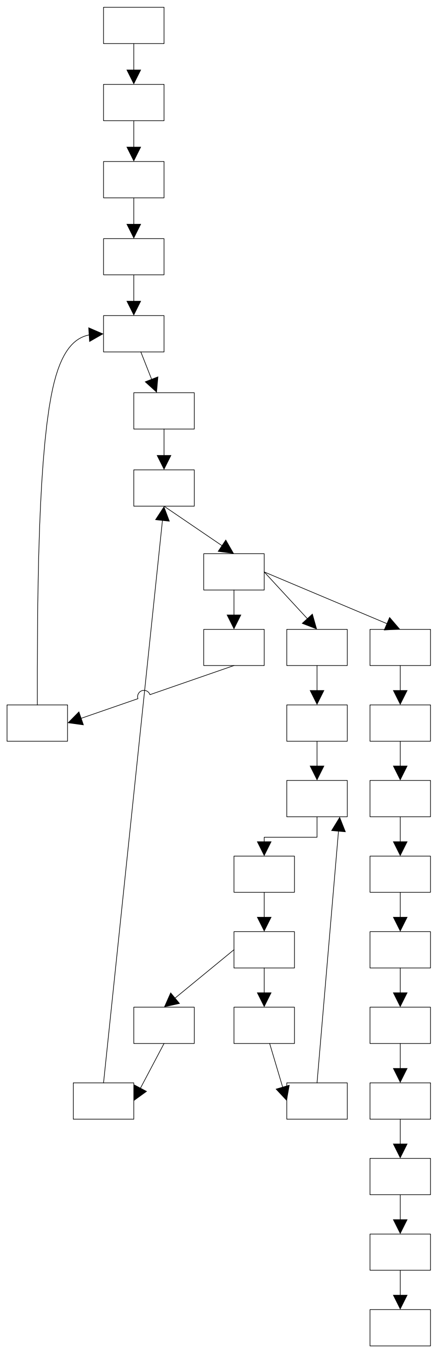


Figure A5.4 Mined heuristic model of digital picture task (high subjective expertise)

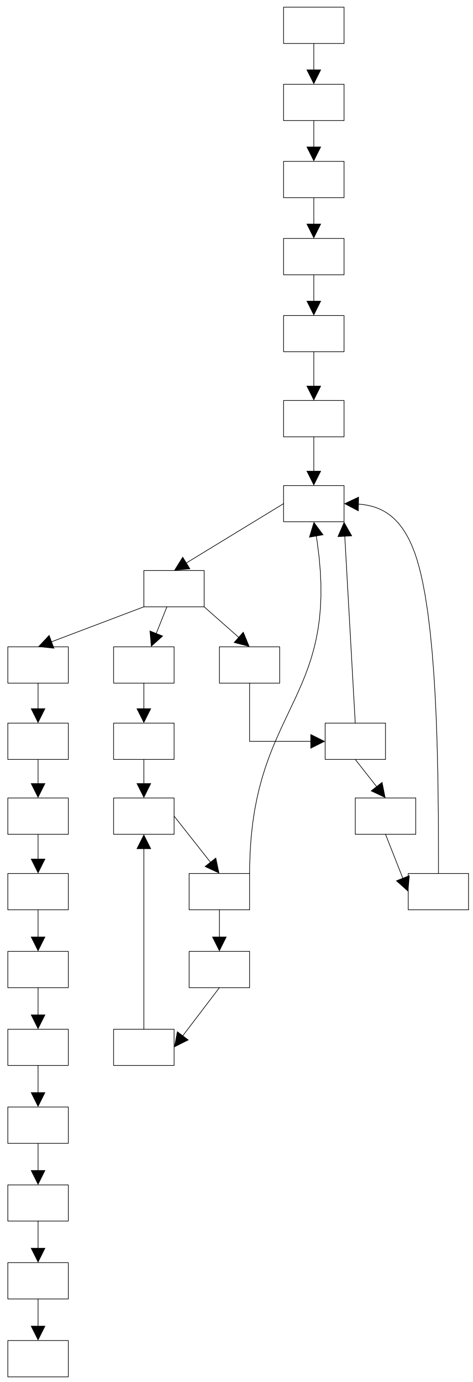


Figure A5.5 Mined heuristic model of digital picture task (low subjective expertise)

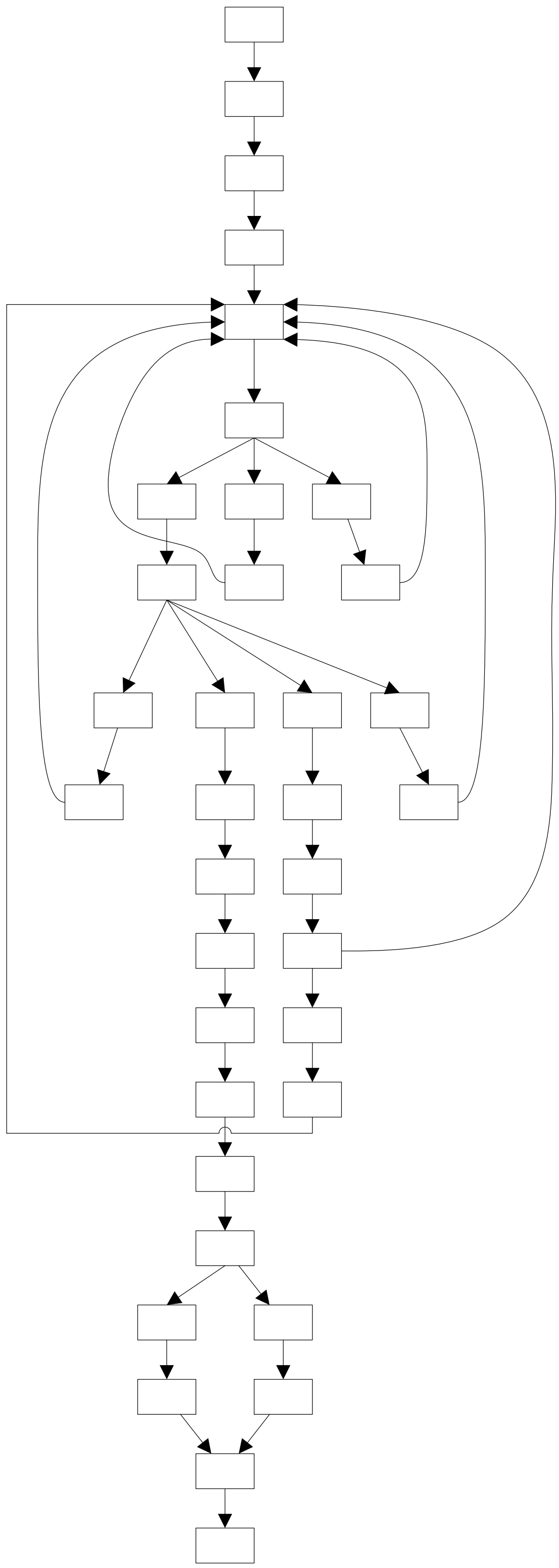


Figure A5.6 Mined heuristic model of switch channel task (high subjective expertise)

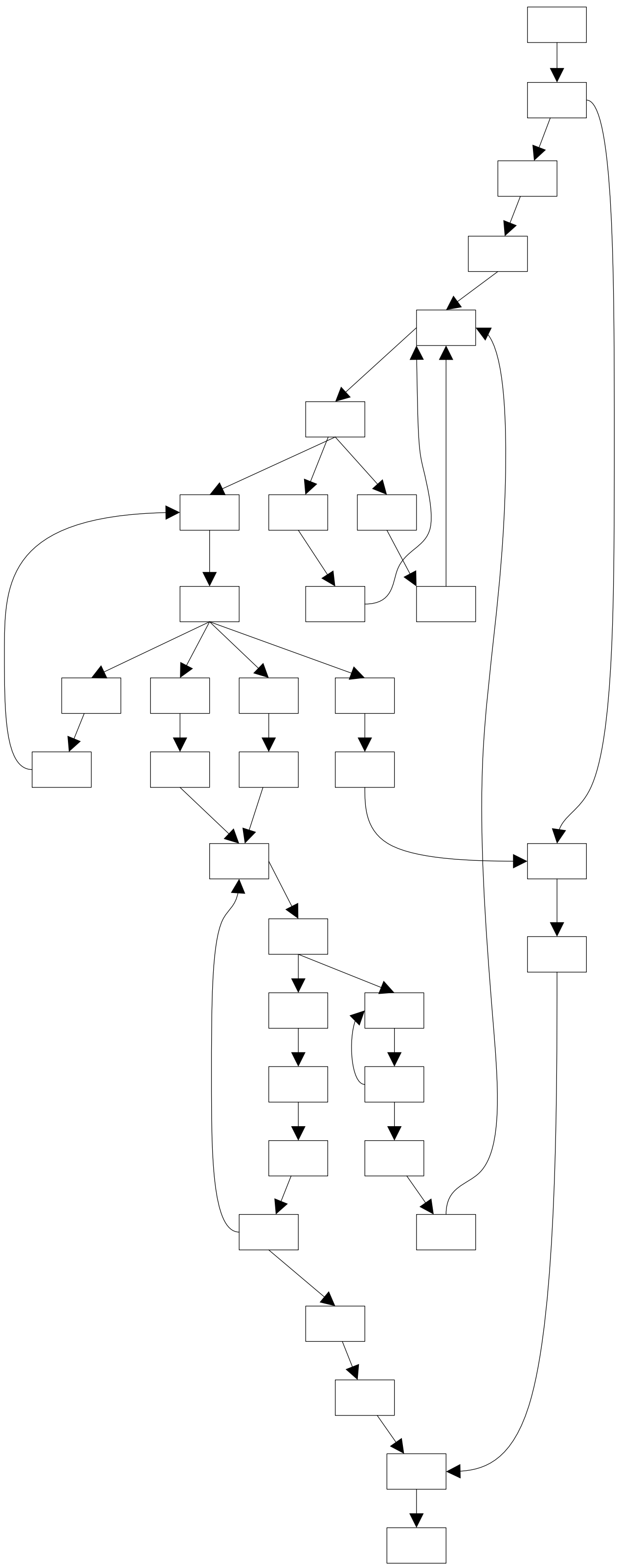


Figure A5.7 Mined heuristic model of switch channel task (low subjective expertise)

Appendix 5.8 Performance sequence diagrams for the tasks

Table A5.6 Performance sequence diagram statistics for digital picture task

	Pattern 0	Pattern 1
Average throughput time (s)	175.89	100.77
Minimum throughput time (s)	51.56	76.52
Maximum throughput time (s)	419.64	153.49
S.D. throughput time (s)	135.68	36.04
Frequency	7	4

Table A5.7 Performance sequence diagram statistics for switch channel task

	Pattern 0
Average throughput time (s)	115.74
Minimum throughput time (s)	91.72
Maximum throughput time (s)	146.58
S.D. throughput time (s)	28.06
Frequency	3

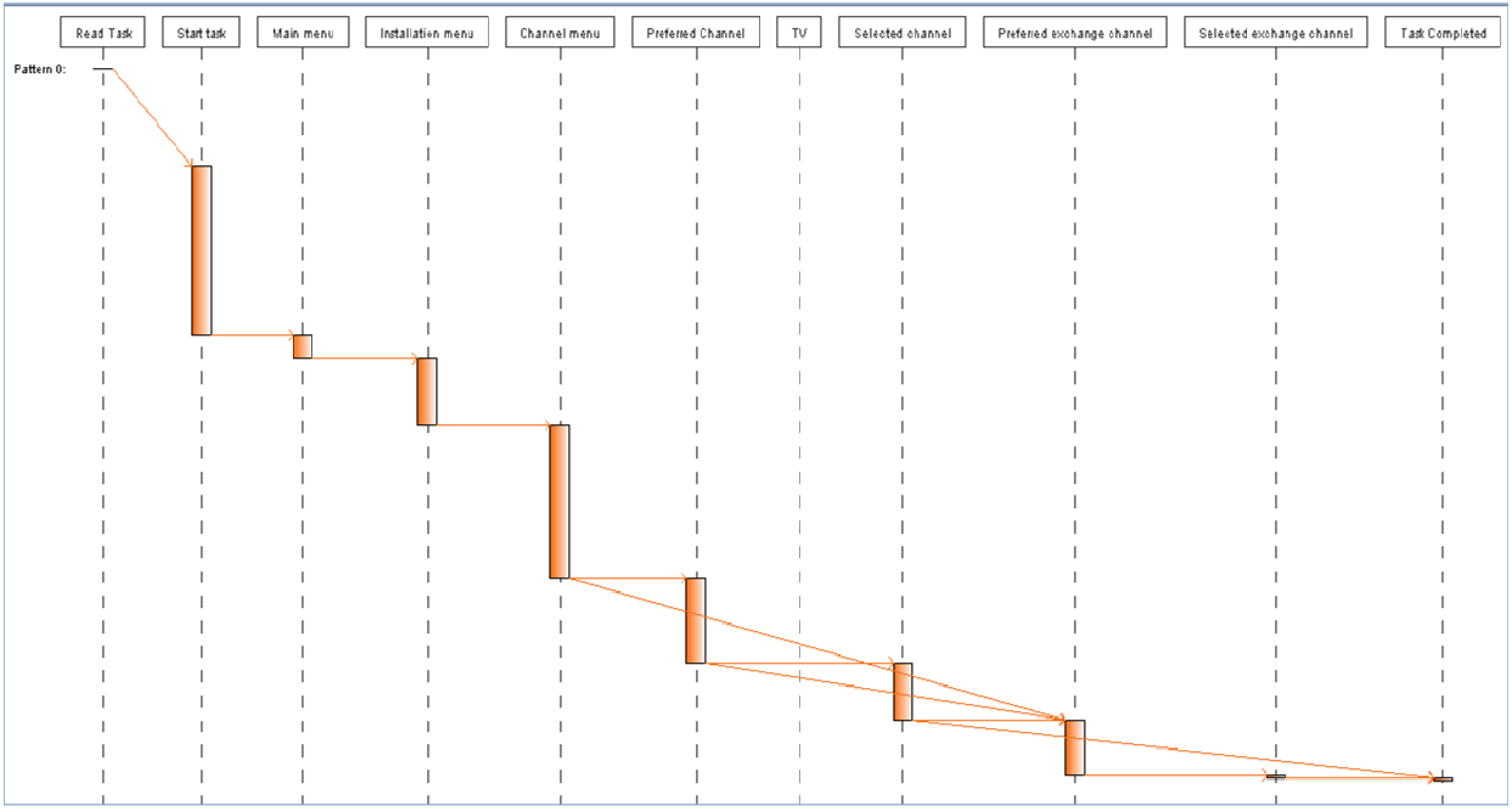


Figure A5.8 Performance sequence diagram for switch channel task

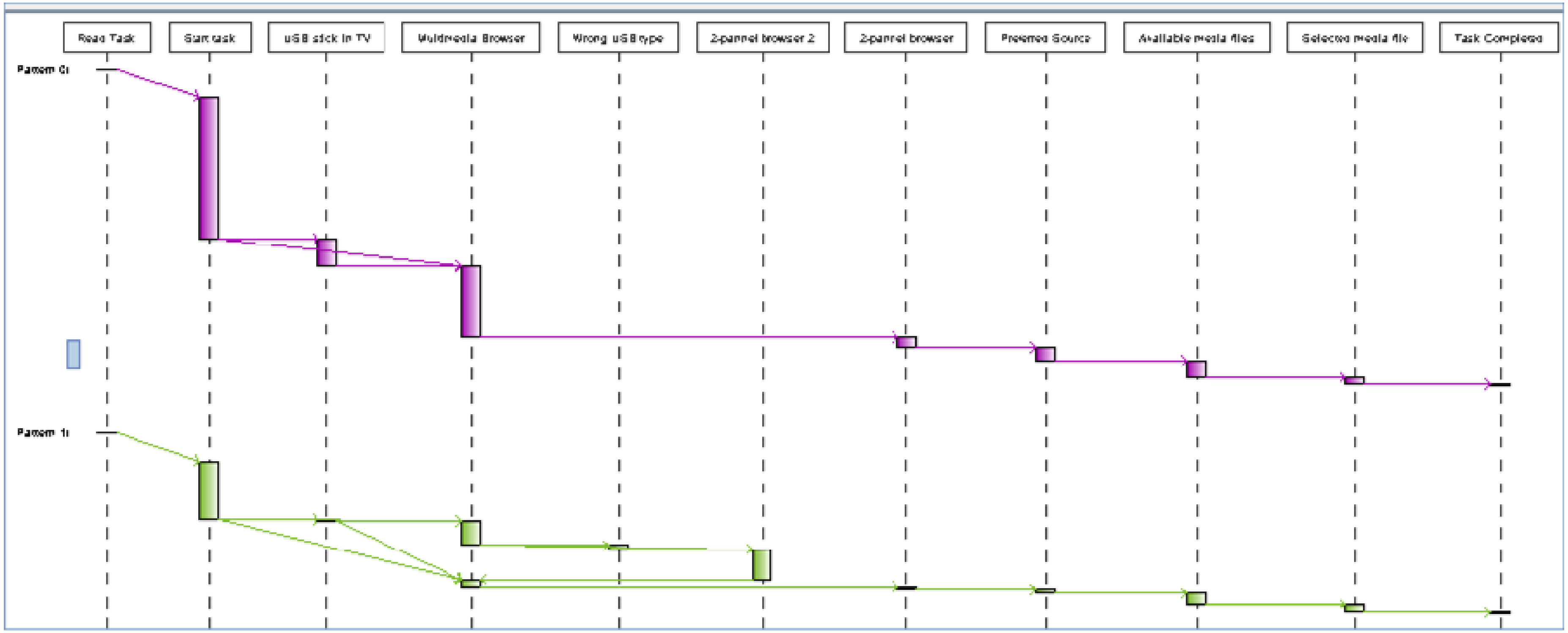


Figure A5.9 Performance sequence diagram for digital picture task

Appendices Chapter 6

Appendix 6.1 Questionnaire items (in Dutch)[31]

Construct	Item	Response scale
Subjective familiarity televisions	1. Ik maak veel gebruik van een televisie. 2. Vergeleken met de meeste andere mensen maak ik weinig gebruik van een televisie	5 point likert scale: • Volledig mee eens • Enigszins mee eens • Niet mee eens, niet mee oneens • Enigszins mee oneens • Volledig mee oneens
Objective familiarity televisions	Hoeveel uur maakt u gemiddeld per week gebruik van een televisie?	Open response
Subjective expertise LCD televisions	1. Vergeleken met de meeste andere mensen weet ik weinig van LCD televisies 2. Ik weet vrij veel van LCD televisies	5 point likert scale
Objective expertise LCD televisions (correct answer printed in **bold**)	1. Waar staat de afkorting LCD voor in de term "LCD televisie"?	Multiple choice: A. Led Coordination Display **B. Liquid Crystal Display** C. Living Colour Display D. Light Compact Display E. Ik weet het niet
	2. Waar staat de term "Hertz" voor als we spreken over een LCD televisie die beelden kan weergeven tot 100 Hertz?	A. Rekeneenheid voor de helderheid van het beeld B. Rekeneenheid voor de contrastwaarde van het beeld **C. Rekeneenheid voor de frequentie van de beeldverversing** D. Rekeneenheid voor de resolutie van het beeld E. Ik weet het niet
	3. Zogenaamde "ruis" of "sneeuw" op het volledige beeldscherm van een LCD televisie wordt meestal veroorzaakt door...?	A. Defecte pixels B. Fout in de software van de televisie **C. Kwaliteit van het inkomende signaal** D. Het beeldformaat E. Ik weet het niet
	4. Wat is het minimum aantal beeldlijnen om volgens de standaard te kunnen spreken van HDTV (High Definition Television)?	A. 480 B. 576 **C. 720** D. 1080 E. Ik weet het niet
	5. Waarom zijn LCD televisies zoveel dunner dan de oudere televisies?	A. De beeldbuis die vroeger heel dik was, is door toepassing van nieuwe technologie in omvang gereduceerd B. LCD televisies bestaan uit veel kleine beeldbuizen C. LCD televisies gebruiken kleine LED lampen in plaats van een beeldbuis **D. LCD televisies gebruiken vloeibare kristallen tussen twee glazen platen met daarachter een grote lamp** E. Ik weet het niet

[31] An (not validated) English translation of the questionnaires in this Appendix can be obtained from the author.

Construct	Item	Response scale
Objective expertise LCD televisions (correct answer printed in **bold**)	6. "Blokkerige" beelden op het scherm van een LCD televisie of velden op het beeldscherm die ineens verspringen of anders gekleurd zijn (zogenoemde MPEG artefacten) kunnen worden veroorzaakt door...?	Check all that apply: A. Defecte pixels **B. Fout in de software van de LCD televisie** **C. Onweer** **D. Compressie van de video** **E. Grote luidsprekerboxen naast de LCD televisie** F. Geen van bovenstaande G. Ik weet het niet
	7. Via welke van de onderstaande aansluitingen op een LCD televisie wordt beeld en/of geluid digitaal doorgegeven?	Check all that apply: **A. HDMI** **B. DVI** C. S-Video D. VGA **E. SCART** F. Geen van bovenstaande G. Ik weet het niet
	8. Op een HD Ready LCD Televisie kunnen analoge televisiesignalen worden weergegeven in HD kwaliteit	True/false statement: • Yes • **No** • Ik weet het niet
	9. De pixels in een LCD televisie worden aangestuurd door software	• **Yes** • No • Ik weet het niet
	10. Rode horizontale of verticale lijnen op het beeldscherm van een LCD televisie worden meestal veroorzaakt door defecte pixels	• Yes • **No** • Ik weet het niet
	11. De beeldkwaliteit van een LCD televisie hangt af van de kwaliteit van de hardware EN van de kwaliteit van de software van de televisie	• **Yes** • No • Ik weet het niet
Product involvement	1. LCD televisies zijn bruikbaar voor mij 2. LCD televisies zijn belangrijk voor mij 3. LCD televisies zijn aantrekkelijk voor mij	5 point likert scale
Product expectations	1. Ik verwacht dat een dergelijke LCD televisie betrouwbaar is 2. Ik verwacht dat een dergelijke LCD televisie een hoge beeldkwaliteit heft 3. Ik verwacht dat een dergelijke LCD televisie foutloos zal werken	5 point likert scale
Picture quality	Hoe vond u de beeldkwaliteit van de LCD televisie voor de zojuist getoonde film?	• Erg slecht • Slecht • Niet goed, niet slecht • Goed • Erg goed
Failure perception	Tijdens het kijken van de film op de LCD televisie, zag u iets wat u zou classificeren als een fout of een storing? Let op: Het gaat hierbij om de film weergegeven op de getoonde televisie	Dichotomous: • Yes • No
Failure description	Geef een korte omschrijving van deze fout / storing.	Open response

Construct	Item	Response scale
Failure impact	Geef op een schaal van 0 tot 5 aan hoe ernstig u deze fout vindt. Een 0 score betekent dat de functie "TV kijken" nog goed bruikbaar is en een 5 score betekent dat de functie "TV kijken" helemaal niet meer bruikbaar is.	Open response in numerical format
Failure experience	Heeft u deze fout zelf al een keer zien optreden in een televisie?	Dichtomous: yes / no
Perceived failure scenario realism	In welke mate vindt u het in de video getoonde scenario met de LCD televisie realistisch?	• Realistisch • Enigszins realistisch • Niet realistisch, niet onrealistisch • Enigszins onrealistisch • Onrealistisch
Failure attribution scenario introduction	In dit onderdeel van de vragenlijst willen wij graag uw mening weten over de kwaliteit van een LCD televisie die wij in de volgende pagina in een opgenomen video laten zien. Deze video laat een LCD televisie zien in een huiskamer waarop een actiefilm wordt getoond die via de kabel worden uitgezonden. Deze LCD televisie is in de winkel verkrijgbaar voor gemiddeld €2500,-. Naast het kijken van televisieprogramma's heeft deze televisie ook de beschikking over functionaliteiten zoals het laten zien van digitale foto's en een PC link waarmee je verbinding kunt maken met een computer. De video duurt ongeveer 90 seconden en wordt afgespeeld zonder geluid. Wanneer u op "Volgende" klikt wordt een nieuwe pagina van de vragenlijst geopend waarin deze video automatisch wordt afgespeeld. Het is hierbij belangrijk om de volledige video te bekijken, van het begin tot het einde.	
Failure attribution (open response)	De volgende vraag is erg belangrijk voor dit onderzoek. Wij willen u vragen bij het beantwoorden van deze vraag voldoende tijd te nemen en zoveel informatie te geven als mogelijk is. Denk aan de oorzaak van de zojuist getoonde fout, gegeven dat deze fout is opgetreden op een echte LCD televisie in een huiskamer. Wie of wat is volgens u verantwoordelijk voor het ontstaan van deze fout? Vul hieronder uw antwoord zo uitgebreid en nauwkeurig mogelijk in.	Open response

Construct	Item	Response scale
Causal dimension scale – locus	1. Deze fout geeft een aspect weer van de kwaliteit van de TV – Deze fout geeft een aspect weer van de kwaliteit van andere dingen 2. Deze fout werd veroorzaakt door een onderdeel van de TV – Deze fout werd veroorzaakt door iets anders 3. Deze fout werd veroorzaakt door iets wat de TV deed – Deze fout werd veroorzaakt door iets wat andere dingen of personen deden	5 point scale with the two items (see items column) as extremes
Causal dimension scale - stability	1. Deze fout zal altijd op deze manier optreden – Het is niet zeker dat deze fout nog een keer op deze manier zal optreden 2. Ik verwacht dat de kwaliteit van deze TV gelijk zal blijven – Ik verwacht dat de kwaliteit van deze TV zal veranderen 3. De oorzaak van deze fout zal nooit veranderen – De oorzaak van deze fout is elke keer verschillend	5 point scale with the two items (see items column) as extremes
Causal dimension scale – controllability	1. Deze fout was te voorkomen door de TV fabrikant of iets of iemand anders – Deze fout was niet te voorkomen door de TV fabrikant of iets of iemand anders 2. Iemand is verantwoordelijk voor deze fout – Niemand is verantwoordelijk voor deze fout	5 point scale with the two items (see items column) as extremes

Appendix 6.2 Validation of objective expertise measurement scale

Table A6.1 Analysis of validity of objective expertise items

Objective expertise item	Mean (= p–value)	S.D.	Point-biserial correlation
1	.62	.487	.570
2	.57	.496	.495
3	.76	.426	.295
4	.21	.411	.355
5	.44	.497	.605
6	.18	.381	.421
7	.54	.499	.489
8	.42	.495	.587
9	.47	.500	.431
10	.25	.434	.403
11	.61	.488	.486

Appendix 6.3 MANOVA results – final solution

Table A6.2 Multivariate tests with objective expertise of televisions and age and product expectations as moderators

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial eta squared	Noncent. Parameter	Observed power
Intercept	Pillai's Trace	.759	359.867	3	343	.000	.759	1079.601	1.000
Age	Pillai's Trace	.023	2,667	3	343	.048	.023	8.000	.649
Failure cause	Pillai's Trace	0.013	1.535	3	343	.205	.013	4.604	.404
Objective expertise group	Pillai's Trace	.062	7.512	3	343	.000	.062	22.537	.986
Failure experience	Pillai's Trace	.031	3.645	3	343	.013	.031	10.934	.796
Failure origin * objective expertise group	Pillai's Trace	.014	1.592	3	343	.191	.014	4.776	.418
Failure origin * failure experience	Pillai's Trace	.003	.400	3	343	.753	.003	1.201	.129
Objective expertise group * failure experience	Pillai's Trace	.005	.606	3	343	.612	.005	1.817	.176
Failure origin * objective expertise group * failure experience	Pillai's Trace	.010	1.186	3	343	.315	.010	3.559	.318

Table A6.3 Test of between-subjects effects for MANOVA with objective expertise and age and TV expectations as moderating variables

Source	Dependent variable	Type III sum of squares	df	Mean square	F	Sig.	Partial eta squared	Observed power[b]
Corrected model	PQ	24.625[a]	8	3.078	3.323	.001	.072	.974
	Impact	5.778[c]	8	.722	.480	.870	.011	.224
	Att_locus	37.103[d]	8	4.638	3.824	.000	.081	.989
Intercept	PQ	168.344	1	168.344	181.727	.000	.345	1.000
	Impact	569.085	1	569.085	378.406	.000	.523	1.000
	Att_locus	357.805	1	357.805	294.978	.000	.461	1.000
Age	PQ	5.565	1	5.565	6.008	.015	.017	.686
	Impact	.010	1	.010	.007	.935	.000	.051
	Att_locus	2.330	1	2.330	1.920	.167	.006	.282
Failure origin	PQ	1.832	1	1.832	1.977	.161	.006	.289
	Impact	1.769	1	1.769	1.176	.279	.003	.191
	Att_locus	2.075	1	2.075	1.711	.192	.005	.257
Objective expertise group	PQ	4.782	1	4.782	5.162	.024	.015	.620
	Impact	.237	1	.237	.158	.692	.000	.068
	Att_locus	23.056	1	23.056	19.007	.000	.052	.992
Failure experience	PQ	6.069	1	6.069	6.551	.011	.019	.723
	Impact	.067	1	.067	.045	.833	.000	.055
	Att_locus	5.244	1	5.244	4.323	.038	.012	.545
Failure origin * Objective expertise group	PQ	.627	1	.627	.677	.411	.002	.130
	Impact	2.404	1	2.404	1.598	.207	.005	.243
	Att_locus	3.998	1	3.998	3.296	.070	.009	.441
Failure origin * failure experience	PQ	.335	1	.335	.362	.548	.001	.092
	Impact	.045	1	.045	.030	.862	.000	.053
	Att_locus	.899	1	.899	.741	.390	.002	.138
Objective expertise group * failure experience	PQ	.403	1	.403	.435	.510	.001	.101
	Impact	.014	1	.014	.009	.924	.000	.051
	Att_locus	1.429	1	1.429	1.178	.279	.003	.191

Source	Dependent variable	Type III sum of squares	df	Mean square	F	Sig.	Partial eta squared	Observed power[b]
Failure origin * objective expertise group * failure experience	PQ	1.495	1	1.495	1.614	.205	.005	.245
	Impact	.424	1	.424	.282	.596	.001	.083
	Att_locus	1.936	1	1.936	1.596	.207	.005	.243
Error	PQ	319.592	345	.926				
	Impact	518.846	345	1.504				
	Att_locus	418.482	345	1.213				
Total	PQ	1571.000	354					
	Impact	6423.000	354					
	Att_locus	4667.000	354					
Corrected total	PQ	344.218	353					
	Impact	524.624	353					
	Att_locus	455.585	353					

a R squared = 0.072 (Adjusted R squared = 0.050)

b Computed using alpha = 0.05

c R squared = 0.011 (Adjusted R squared = -0.012)

d R squared = 0.081 (Adjusted R squared = 0.060)

Appendices Chapter 7

Appendix 7.1 Added objective expertise items (in Dutch)[32]

Construct	Item	Response scale
Additional items to measure objective expertise of LCD televisions (correct answer printed in **bold**)	12. De automatische installatie van televisie zenders heeft de zender RTL4 op uw LCD televisie op kanaal 20 gezet. U wilt dit veranderen en RTL4 op kanaal 4 installeren. Hoe pakt u dit aan?	Multiple choice: A. Ik zet de televisie uit en vervolgens weer aan B. Ik doe de automatische installatie van de televisie zenders opnieuw **C. Ik ga in het menu van de televisie naar handmatig instellen van de zender** D. Ik zet in het menu van de televisie de televisie terug naar de fabrieksinstellingen E. Ik weet het niet
	13. Stelt u zich de volgende situatie voor: U heeft een LCD televisie die is aangesloten op een analoog kabelsignaal. Verder heeft u geen randapparatuur zoals een set-top box of harddisk recorder aangesloten. Na enkele maanden van gebruik doet zicht het volgende probleem voor: Nadat u de LCD televisie heeft aangezet, merkt u op dat alle zenders op uw televisie geen beeld en geen geluid hebben terwijl het power lampje van de televisie aan is. Wat is het eerste dat u doet om het probleem op te lossen?	Multiple choice: A. Ik installeer de zenders opnieuw **B. Ik controleer de kabelaansluiting van het televisie signaal** C. Ik zet de televisie uit en daarna weer aan D. Ik pas de instellingen voor kleur, contrast, tint en helderheid aan E. Ik weet het niet
	14. Stelt u zich de volgende situatie voor: U heeft een LCD televisie gekocht en u heeft deze thuis aangesloten op een analoog kabelsignaal. Verder heeft u geen randapparatuur zoals een set-top box of harddisk recorder aangesloten. Na installatie van alle televisie zenders merkt u op dat op alle zenders het geluid goed is, maar een slechte kleur of soms zelfs geen beeld hebben. Wat is het eerste dat u doet om het probleem op te lossen?	Multiple choice: **A. Ik pas de instellingen voor kleur, contrast, tint en helderheid van het beeld aan** B. Ik controleer of de stekker goed in het stopcontact zit C. Ik installeer de televisie zenders opnieuw D. Ik controleer of de kabel voor het televisiesignaal goed is aangesloten E. Ik weet het niet

[32] An (not validated) English translation of the questionnaires in this Appendix can be obtained from the author.

Appendix 7.2 Experiment questionnaire items (in Dutch)

Construct	Item	Response scale
Experiment introduction	Dit experiment gaat over de kwaliteit van LCD televisies. De resultaten van dit onderzoek worden gebruikt om de kwaliteit van toekomstige LCD televisies te verbeteren. Voor dit experiment stellen wij jou in deze vragenlijst tien vragen over een video die getoond wordt op de LCD televisie die voor je staat.	
Product involvement	See Appendix 6.1	
Product expectations	See Appendix 6.1	
Introduction to scenario with failure	Op een in de winkel verkrijgbare LCD televisie is, na een half jaar gebruik zonder problemen, tijdens het kijken van TV programma's via de kabel een probleem geconstateerd met de kwaliteit van het beeld. Deze LCD televisie was alleen aangesloten met een coax kabel op een analoog kabelsignaal en op het stroomnet. Verder was de TV niet aangesloten op andere randapparatuur zoals een DVD speler of set-top box. Dit probleem hebben experts op het gebied van TV ontwikkeling in juni 2008 met professionele apparatuur vastgelegd in een video die via de "monitor-out" uitgang van de LCD televisie op een computer is opgenomen. Deze opname heeft dus geen invloed op de beeldkwaliteit: de video laat zien wat er op dat moment op het beeldscherm van de LCD televisie te zien was. Het geconstateerde probleem met de beeldkwaliteit kan dus ook in elke vergelijkbare thuissituatie voorkomen. In dit experiment zijn wij geïnteresseerd in jouw mening over dit probleem met de beeldkwaliteit. Zometeen ga je op de LCD televisie kijken naar de video waarin dit probleem is vastgelegd. De video duurt in totaal twee minuten. Vervolgens worden er in de volgende pagina's van deze vragenlijst een aantal vragen over deze video gesteld. Het is hierbij belangrijk om de volledige video te bekijken voordat je verder gaat met de vragenlijst. Na het bekijken van de video kun je verder gaan met de vragen over deze video door op "Volgende" te klikken. Vraag nu aan de observator om de video te starten.	
Picture quality	See Appendix 6.1	
Failure description	See Appendix 6.1	
Failure impact	See Appendix 6.1	
Failure experience	See Appendix 6.1	

Construct	Item	Response scale
Failure attribution (open response)	De volgende vraag is erg belangrijk voor dit onderzoek. Wij willen je daarom vragen bij het beantwoorden van deze vraag voldoende tijd te nemen en zoveel informatie te geven als mogelijk is. Denk aan de oorzaak of oorzaken van de zojuist getoonde fout, gegeven dat deze fout is opgetreden op een echte LCD televisie in de beschreven situatie: "Deze LCD televisie was alleen aangesloten met een COAX kabel op een analoog kabelsignaal en op het stroomnet. Verder was de TV niet aangesloten op andere randapparatuur zoals een DVD speler of set-top box". Wat zijn volgens jou alle mogelijke, realistische oorzaken van dit probleem? Ook al ben je geen expert op het gebied van televisies, probeer hier zo volledig mogelijk te beschrijven wat jij denkt over alle mogelijke oorzaken van dit probleem. Vul hieronder jouw antwoord puntsgewijs zo volledig en nauwkeurig mogelijk in.	Open response
Causal dimension scale – locus	See Appendix 6.1	
Causal dimension scale – stability	See Appendix 6.1	
Causal dimension scale – controllability	See Appendix 6.1	
Failure attribution (multiple choice)	Wij willen je nu nog een laatste keer vragen om na te denken over de oorzaak van het probleem met de beeldkwaliteit. Hieronder staan een aantal categorieën van mogelijke oorzaken voor het probleem met de beeldkwaliteit van de LCD televisie in de zojuist getoond video. Kruis aan welke categorieën volgens jou realistische oorzaken kunnen zijn van het probleem.	Select all that apply: A. Fout in de hardware van de televisie B. Slechte ontvangst van het analoge kabelsignaal C. De gebruiker heeft de televisie verkeerd ingesteld D. Slechte kwaliteit van de opname door de televisiezender die het programma uitzendt E. Fout in de software van de televisie F. Storing in de omgeving van de televisie G. Ik weet het niet
Problem solving strategy	Als jij dit probleem thuis zou ondervinden met deze LCD televisie, wat zou jij als eerste doen om dit probleem op te lossen?	Open response
Perceived failure scenario realism	See appendix 6.1	

Appendix 7.3 Validation of objective expertise measurement scale

Table A7.1 Analysis of validity of all objective expertise items

Objective expertise item	Mean (= p–value)	S.D.	Point-biserial correlation
1	.90	.307	.366
2	.97	.184	.354
3	.86	.348	.113
4	.31	.467	.240
5	.88	.329	.177
6	.34	.479	.204
7	.48	.504	.357
8	.72	.451	.558
9	.79	.409	.426
10	.53	.503	.068
11	.78	.421	.697
12	.98	.131	-.150
13	.74	.442	-.003
14	.17	.381	-.081

Table A7.2 Analysis of validity of included objective expertise items

Objective expertise item	Mean (= p–value)	S.D.	Point-biserial correlation
1	.90	.307	.419
3	.86	.348	.197
4	.31	.467	.245
5	.88	.329	.143
6	.34	.479	.284
7	.48	.504	.440
8	.72	.451	.560
9	.79	.409	.528
11	.78	.421	.660

Appendix 7.4 MANOVA results for the final solution

Table A7.3 Multivariate tests with objective expertise and failure impact as independent variables and age as Moderating variable

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial eta squared	Noncent. Parameter	Observed power
Intercept	Pillai's Trace	0.801	68.397	3	51	.000	.801	205.190	1.000
Age	Pillai's Trace	0.037	0.648	3	51	.588	.037	1.944	0.176
Failure impact	Pillai's Trace	0.079	1.465	3	51	.235	.079	4.395	0.365
Objective expertise group	Pillai's Trace	0.124	2.407	3	51	.078	.124	7.222	0.568
Failure impact * objective expertise group	Pillai's Trace	0.035	0.610	3	51	.611	.035	1.831	0.168

Table A7.4 Test of between-subjects effects for MANOVA with objective expertise and failure impact as independent variables and age as moderating variable

Source	Dependent variable	Type III sum of squares	df	Mean square	F	Sig.	Partial eta squared	Observed power[b]
Corrected model	PQ	4.170[a]	4	1.043	1.439	.234	.098	.416
	Impact	337.196[c]	4	84.299	.984	.424	.069	.290
	Att_locus	.358[d]	4	.090	1.180	.330	.082	.345
Intercept	PQ	15.120	1	15.120	20.870	.000	.283	.994
	Impact	1762.552	1	1762.552	20.582	.000	.280	.994
	Att_locus	11.221	1	11.221	147.736	.000	.736	1.000
Age	PQ	.001	1	.001	.002	.966	.000	.050
	Impact	116.691	1	116.691	1.363	.248	.025	.209
	Att_locus	.040	1	.040	.526	.471	.010	.110
Failure impact	PQ	2.707	1	2.707	3.736	.059	.066	.475
	Impact	12.251	1	12.251	.143	.707	.003	.066
	Att_locus	.008	1	.008	.099	.754	.002	.061
Objective expertise group	PQ	1.280	1	1.280	1.767	.189	.032	.257
	Impact	64.459	1	64.459	.753	.390	.014	.136
	Att_locus	.316	1	.316	4.160	.046	.073	.517
Failure impact * Objective expertise group	PQ	.016	1	.016	.022	.882	.000	.052
	Impact	154.737	1	154.737	1.807	.185	.033	.262
	Att_locus	.004	1	.004	.049	.826	.001	.055
Error	PQ	38.399	53	.725				
	Impact	4538.683	53	85.636				
	Att_locus	4.026	53	.076				
Total	PQ	255.000	58					
	Impact	42897.000	58					
	Att_locus	142.667	58					
Corrected total	PQ	42.569	57					
	Impact	4875.879	57					
	Att_locus	4.384	57					

a R squared = .098 (Adjusted R squared = .030)

b Computed using alpha = .05

c R squared = .069 (Adjusted R squared = -.001)

d R squared = .082 (Adjusted R squared = .012)

Curriculum Vitae

Jeroen Keijzers was born in Roosendaal, the Netherlands, on December 18th, 1981. In 2005 he received his Masters degree (cum laude) in Industrial Engineering and Management Science from Eindhoven University of Technology. The topic of his graduation project, performed at Philips Applied Technologies, was the set-up of consumer test strategies to identify user-perceived failures in innovative consumer electronics.

In November 2005, he started his Ph.D. research project at the sub department of Quality and Reliability Engineering at the faculty of Technology Management at Eindhoven University of Technology. This research project was continued from January 2008 onwards at the sub department of Business Process Design at the faculty of Industrial Design at Eindhoven University of Technology. It was performed in cooperation with various academic partners (Embedded Systems Institute, Design Technology Institute, Delft University of Technology, Leiden University and University of Twente) and industrial partners (NXP Semiconductors, TASS, Philips Consumer Electronics and IMEC).

Since November 2009 he is working as a researcher/lecturer at the sub department of Business Process Design at the faculty Industrial Design at Eindhoven University of Technology. His research interests include analyzing the consumer's perception of product value and the design of value creation models with specific focus on the analysis and modeling of stakeholder needs in designing intelligent systems in an open innovation context. Besides research he is also involved in both coaching students and lecturing several courses related to business process design.