Improving maintenance-sensitive products using online product reviews

the effect of adding product features on consumer satisfaction through consumers' perceptions of capability, usability and maintainability

van Kemenade, R.

Award date:
2018

Eindhoven, February 2018

Improving Maintenance-Sensitive Products using Online Product Reviews

The effect of adding product features on consumer satisfaction through consumers' perceptions of capability, usability and maintainability

by
R. (Robin) van Kemenade

BSc Natuurwetenschap & Innovatiemanagement – Utrecht University
Student identity number 0649945

in partial fulfilment of the requirements for the degree of

Master of Science
in Innovation Management

Supervisors:

prof. dr. E.J. (Ed) Nijssen, TU/e, ITEM
dr. S. (Shantanu) Mullick, TU/e, ITEM
dr. S.E.C. (Sarah) Gelper, TU/e, ITEM

Company Supervisor:

F. (Frank) Meinster MSc, Philips Benelux, Personal Health

TUE. School of Industrial Engineering.
Series Master Theses Innovation Management

Subject headings: online product reviews, new product development, feature fatigue, maintainability

Acknowledgements

After obtaining my bachelor's degree in Science and Innovation Management at Utrecht University, I was determined to choose a master's that shared the orientation of my bachelor's, innovation, and provided opportunities to experience innovation in practice. My expectations of the master Innovation Management at Eindhoven University of Technology were high and by now, I can proudly say that these are completely met. I appreciate every single course, project, lecture, exam, teamwork session, professor, and fellow student that have enabled me to achieve my full potential and become the almost-graduated student I am today. Especially the final part, this master thesis, has not been easy. Fortunately, many people have been so supportive to help me overcome the issues and submit a thesis I am proud of. Therefore, I would like to thank everyone that has been involved in my graduation process.

Especially, I would like to thank my first supervisor prof. dr. Ed Nijssen for providing guidance to my master thesis project. Where I sometimes lost focus or motivation, you always managed to get me back on track and provide me with helpful feedback, regardless of your busy schedule. Furthermore, I am very grateful to dr. Shantanu Mullick. Thank you for helping me solve complex methodological issues as well as reviewing my work multiple times. It has been an honour to work with these two professionals. Also, I would like to thank dr. Sarah Gelper for being the third assessor of my thesis.

My graduation internship at Philips has been an unforgettable experience. It not only enabled me to perform the case study of my thesis; it has also brought me a lot of knowledge, professional skills, and fun. I would like to thank my company supervisor Frank Meinster for making me feel at home from day one, for the brainstorm sessions, and crucial insights into Philips' full automatic espresso machines as well as their consumers. Additionally, I really appreciate the support and good fellowship of the coffee team and other interns I have worked with.

I would furthermore like express my appreciation to all others that have been supportive and understanding during my thesis period: my family, friends, fellow students, colleagues and even neighbours. Without you, I could not have come this far. Particularly, I want to thank my study buddy and dear friend Brent Verschoor for being the key reviewer of my thesis. Furthermore, many gratitude goes to all FAEM users that helped me to validate the data pre-processing steps. Lorain Giesbers, thanks for providing valuable and motivating feedback on my concept thesis. Last, but certainly not least, I would like to praise Bram van der Sommen for his unconditional support, motivational talks and feedback. Thank you so much!

Graduating will be a dream come true. At the same time, it marks the end of the most important part of my life so far. In addition to the study-related skills and knowledge obtained over the past years, I have made big steps forward in terms of personal, social and professional development. I however believe that learning has only just begun, and am eager to never stop learning. I am truly looking forward to bringing the great amounts of knowledge I have obtained during my studies into practice, to initiate successful and meaningful innovation.

Management Summary

The generation and use of online product reviews has increased drastically over the past decades. Using online product reviews to identify product improvements has proved relevant to product researchers as well as developers. Online reviews are consumers' product evaluations arising from actual purchase and usage situations. Reviews are therefore found to be more reliable, authentic and better representative of the voice of the consumer than traditional consumer research methods, such as conjoint analysis. This study focuses specifically on extracting insights for new product development from product reviews posted on designated product pages.

To capture the voice of the consumer, online review analysis identifies consumer needs based on the topics that consumers mention in their reviews, which are referred to as consumer-defined product attributes. As opposed to previous online review research, this study also accounts for manufacturer-defined product features to represent the complexity of products. Increased complexity of a product, arising from popular, feature-based differentiation strategies, can have a two-sided effect on consumer satisfaction: through increased perceived capability and decreased perceived usability of the product. This two-sided effect is defined as feature fatigue, and measured by Thompson et al. (2005) based on simulated rather than real usage situations. This current study enriches the findings of Thompson et al. (2005) by operationalising capability and usability as bundles of product attributes extracted from online reviews. Furthermore, an additional dimension of consumer needs is assessed, maintainability, applying to products that require frequent or extensive maintenance by the user to keep the product in working condition. The research question addressed is "*To what extent do perceived capability, usability, and maintainability of maintenance-sensitive products influence the effect of adding features on consumer satisfaction?"*. Accordingly, the conceptual model tested in this study connects features with consumer satisfaction through mediation of capability, usability, and maintainability.

The case study used to measure these effects concerns Philips' full automatic espresso machines in the Netherlands. Philips does acknowledge increasing consumer needs for maintainability of these machines, but lacks insights into how these needs relate to capability and usability needs. Furthermore, as these machines become increasingly complex due to added features, it is relevant to measure how consumer needs for capability, usability and maintainability change as the number of features increases. Therefore, online reviews were collected and analysed of three full automatic espresso machines of Philips, differing in number of features. Extracting consumer needs from online product reviews involves of identifying consumer-defined product attributes, extracting consumers' opinions towards these attributes and scoring these opinions based on sentiment analysis. In this particular study, product attributes and scores were grouped into either capability, usability and maintainability to represent underlying consumer needs. With satisfaction represented by consumers' overall product ratings in reviews, the effects of adding product features were measured using path analysis.

The results of the path analysis show a negative, significant partial effect of features on maintainability, indicating that perceived maintainability decreases when features are added. Furthermore, a positive, significant effect of maintainability on satisfaction was found. The combined indirect effect of maintainability on the features-satisfaction relationship was not significant when simultaneously accounting for the significant, positive mediating effects of capability and usability. However, it is still recommended to product developers of Philips' full automatic espresso machines to account for maintainability as it is currently perceived as low to even negative for the types of machines studied.

While the results of this study are product-specific, generalisations can be drawn for other industries of maintenance-sensitive products. The findings of this study demonstrate that maintainability is an

important factor of consumer satisfaction, and becomes problematic when the number of product features increases. Product developers of maintenance-sensitive products are therefore recommended to account for maintainability as design requirement, especially when products become increasingly technologically advanced. However, needs for maintainability should be balanced against needs for capability and usability. In absence of priority information, product developers may address the wrong development objectives or even ease-of-development, risking new product's consumer satisfaction and market success.

It should be noted that maintainability might start to play a more prominent role after the machine has been in use for a longer period. Several maintenance actions are not required until multiple months of usage have passed, depending on different factors, such as usage frequency. Unfortunately, this study was unable to capture consumers' long-term-based product evaluations since the average time between purchase and posting a review was 1.38 months. Other limitations of this research include the scalability of the approach, due to manual pre-processing steps, and the measure used for consumer satisfaction, which might be affected by consumers' different interpretations of the star-rating scale. These limitations of the study provide relevant directions for future research, including the measurements of the studied effects over time, and comparing or developing data pre-processing tools as well as measures for consumer satisfaction based on online product reviews.

Table of Contents

1. Introduction

Along with the accelerating growth of the internet and digital applications, both generation and usage of online product reviews have increased exponentially over the past decade. Consumers increasingly base their purchase decisions on online reviews by other consumers (Hardey, 2011), and 81% indicates to receive helpful advice from online reviews (Xiao et al, 2016). More and more firms nowadays acknowledge the relevance of this source of online word-of-mouth to various business practices. Specifically, online reviews have been shown to effectively describe consumer needs, which can be translated into important insights for a firm's new product development (NPD) processes (e.g., Decker & Trusov, 2010; Wang & Wang, 2014; Qi et al., 2016). The wealth of information embedded in online reviews provides valuable opportunities to improve new products and create competitive advantage.

Online reviews are consumers' product evaluations arising from real usage situations, which makes them more suitable to identifying consumer needs than traditional consumer research methods, such as conjoint analysis and focus groups. Even before the first review sites were launched, e.g., Epinions in 1999 (Rivlin, 2005), Srinivasan et al. (1997) suggested that consumer needs can be more accurately predicted using product-in-use research. Accordingly, reviews were later shown to be more authentic, reliable and representative of the voice of the consumer (VOC) than research methods based on hypothetical or simulated usage situations (Bickart & Schindler, 2001; Ghose & Ipeirotis, 2007). This study emphasises the use of online reviews in consumer research for NPD, focussing specifically on reviews that are posted on designated product pages, rather than on social media, blogs or vlogs.

Online review research generally identifies consumer-defined product attributes from reviews, enabling detection of both manifest and latent consumer needs (Archak et al., 2011). In contrast, the more established consumer research method conjoint analysis defines products as a bundle of manufacturer-predefined features (Goldenberg et al., 2003; Mukherjee & Hoyer, 2001). In the absence of the VOC, consumer research might fail to capture what actually satisfies consumers (Thompson et al., 2005). To denote the difference between attributes and features as conceptualised in this study, the following definitions are used. *Product attributes* are subjective characteristics of a product that are endogenously perceived and evaluated by consumers, expressed in their product evaluations arising from usage situations, i.e., costs and benefits of a product to an individual consumer (Abraham-Murali & Littrell, 1995; Thompson et al., 2005; Kangale et al., 2016). Examples of product attributes are perceived aesthetics, personalisation, and ease-of-use. *Product features* are objective characteristics of a product that are exogenously defined by manufacturers, expressed in product descriptions in (online) shops, i.e., mostly technological opportunities of a product that distinguishes it from competitors (Kotler & Armstrong, 2007; Archak et al., 2011; Thompson et al., 2005; Netzer et al., 2008). Examples of product features are type of display, operation speed, size, and battery life. In online review analysis, product attributes are the topics that consumers refer to in their product reviews (Decker & Trusov, 2010). Online review research generally refrains from using product features as source for consumer need identification, since usage-related needs might be overlooked, resulting in a NPD solution that does not maximise consumer satisfaction. However, since many firms differentiate their products based on additional, unique features (Goldenberg et al., 2003), it would be relevant to consider the effect of adding features on consumer satisfaction, accounting for consumers' evaluations of product attributes that arise from using the feature-rich product.

Popular, feature-based differentiation strategies assume that each positively valued feature increases consumer satisfaction and new product success (Nowlis & Simonson, 1996; Thompson et al., 2005). An often-overlooked issue is, however, that the cost arising from the complexity to use a feature-rich product often exceeds the benefit of having many functions (Burke, 2013; Mick & Fournier, 1998).

Thompson et al. (2005) referred to this phenomenon as feature fatigue, and examined its effects on consumer satisfaction. Higher levels of features increase satisfaction through increased *perceived capability*, i.e., consumers' perceptions of the product's ability to perform the desired functions. But at the same time, adding features may decrease satisfaction througsdecreased *perceived usability*, i.e., consumers' perceptions of the ease of learning and using the desired functions (Thompson et al., 2005). This implies that adding product features could in fact harm consumer satisfaction and new product success. As technology advances, product developers should thus be aware of consumer needs for capability and usability for different levels of features. Consumer needs for capability- and usability-related product attributes can effectively be extracted from online product reviews to examine their role in the features-satisfaction relationship (et al., 2014).

Besides capability and usability, specific product types are associated with a third category of consumer needs that should gain attention in NPD: maintainability (Slavila et al., 2005; Pecht, 2009). Maintainability, in this context, applies to products that require regular maintenance by the user, rather than maintenance as a service by a firm, including types of software (Slavila et al., 2005; Mari, 2003), clothes (Seva et al., 2007), copy machines (Shimomura et al., 1995), washing machines and dryers (Seliger et al., 1994), kitchen appliances (Lok et al., 2013) and coffee machines (Frizziero, 2014). Consumers may expect maintenance operations (e.g., replacing an ink cartridge) to be performed quickly, easily and with no training (Hooks & Farry, 2001, p.76). Accordingly, this study defines *perceived maintainability* as consumers' perceptions of the ease of cleaning and maintaining the product to keep the desired functions in working condition (Seva et al., 2007). Maintainability concerns the entire life cycle of the product and therefore, it is crucial to integrate maintainability aspects in the early product development stage (Mari, 2003; Slavila et al., 2005). NPD may otherwise fail to address specific, latent, consumer needs arising from maintaining the product. Instead, product developers may prioritise capability, usability, or even ease of development (Hooks & Farry, 2001, p.153), risking consumer satisfaction and new product success.

The role of maintainability has been acknowledged in previous research, particularly, software and digital products literature (e.g., Boehm et al., 1976; Kekre et al., 1995; Slavila et al., 2005; Pecht, 2009). In the software area the importance of maintainability to consumers' perceived product quality and satisfaction has been demonstrated, since maintainability directly affects availability and reliability of the product to the consumer (Mari, 2003). Required maintenance should be easy, time-efficient to perform, and components should be easy to access to consistently deliver the desired performance (Pecht, 2009). In the context of clothing, Seva et al. (2007) demonstrated that products that are easily maintained elicit positive emotional reactions from consumers since they imply a convenient and stress-free lifestyle. These findings indicate that, for maintenance-sensitive physical products, perceived maintainability can also be regarded to as factor of consumer satisfaction, in addition to capability and usability. Furthermore, as the degree of maintainability is often linked directly to the complexity of products (e.g., Slavila et al., 2005; Pecht, 2009), it is assumed that maintainability becomes more of an issue when complexity of a product, i.e., the number of features, increases. However, no prior literature was found to explicitly include maintainability as predictor of consumer satisfaction of physical consumer products, nor in relation to increasing product features. This study aims to close this gap by integrating maintainability in the conceptual model of Thompson et al. (2005). The effect of adding product features on consumer satisfaction is measured, accounting for consumers' perceptions of capability, usability and maintainability extracted from online reviews.

One industry of feature-rich products in which both firms and consumers increasingly adhere to maintainability, on top of capability and usability, is the full automatic espresso machine (FAEM) industry. FAEMs have shown initial growth in Western European coffee machine market over the past

years and have the potential to capture a great market share in the near future (Morris, 2013; Businesswire, 2016). FAEMs are increasingly featured with advanced operating, grinding, brewing, and milk frothing technologies. Compared to drip filter and single-serve coffee machines, extensive and frequent maintenance is required by the user to keep a FAEM working properly (Philips, 2017). One of the main players in the FAEM category, Philips, recognises consumer needs for maintainability. The problem that Philips faces is, however, that it lacks knowledge of consumers' perceptions of capability, usability and maintainability for different levels of product features and their relative effects on consumer satisfaction (F. Meinster, personal communication, March 6, 2017). This knowledge is considered crucial to Philips' objective to once again obtain a dominant design position (Christensen et al., 1998) in the coffee machine market, as did their Senseo in the 2000s (Griffioen, 2008; Van Kralingen, 2007). This research is therefore performed based on the case study of Philips' FAEMs. As consumers' coffee preferences and associated needs significantly differ across countries and cultures (Morris, 2013), consumer insights should be obtained and applied region-specifically. This research focuses on the FAEM market in the Netherlands.

This study proposes a new, integrated model to measure consumer satisfaction based on perceived capability, usability and maintainability along a product range with increasing levels of features. The research question that is addressed in this study is: "*To what extent do perceived capability, usability, and maintainability of maintenance-sensitive products influence the effect of adding features on consumer satisfaction?".* To account for increasing product features, three different types of Philips' FAEMs were assessed, each having a different number of features (Thompson et al., 2005). Capability, usability, and maintainability are represented by clustered consumer evaluations of product attributes identified from reviews. Consumers' star ratings in reviews are used as measure of consumer satisfaction (Decker & Trusov, 2010).

This research does not only provide relevant managerial implications for improving Philips' FAEMs and other firms' maintenance-sensitive product types. It furthermore connects and contributes to several academic research disciplines, including NPD, design, innovation, marketing, management, and online product review literature. The contribution to these disciplines is twofold. Firstly, the relevance of maintainability to consumer satisfaction is demonstrated in relation to the trend of increasing product features. Secondly, this study expands the range of opportunities for online reviews as data source for consumer research.

The theoretical framework underlying this research is described in Chapter 2, combining relevant literature insights regarding online product reviews, feature fatigue, perceived capability, usability, and maintainability. A conceptual model is developed that hypothesises the relationships required to answer the research question of this study. The methodology is discussed in Chapter 3. Measuring consumer needs from online product reviews requires several methodological steps, including the collection of reviews in a database, the identification of product attributes, the categorisation of attributes into capability, usability and maintainability, the extraction of consumer opinions towards product attributes, and sentiment analysis. Subsequently, mediation analysis is conducted to test the conceptual model, and path analysis is applied to measure the relationships between number of features, attribute categories and consumer satisfaction. Chapter 4 reports the results of this study. Chapter 5 includes the discussion, interpreting the findings and answering the research question. Theoretical contributions, managerial implications, limitations and direction for future research are discussed in the conclusion in Chapter 5.

2. Theoretical Background

As this study is based on the analysis of online product reviews, it discusses relevant previous online product review literature first. Section 2.1 introduces the concept of online product reviews by describing characteristics as well as various applications of online product reviews. Section 2.2 narrows the theoretical lens to the scope of this study, focussing on previous online product research that is used for NPD. To integrate the voice of the consumer in NPD, product attributes, opinions about these attributes and the sentiment polarity of these opinions are generally extracted from online product reviews. To represent the feature fatigue effect, this study clusters the product attributes and scores into underlying consumer needs for capability, usability, and maintainability. The effect of adding product features on consumer satisfaction is described in section 2.3, including the feature fatigue effect through capability and usability (2.3.1), and the introduction of maintainability as additional determinant of the feature-satisfaction relationship (2.3.2). The theoretical background closes with the development of the conceptual model and hypotheses in section 2.4.

First of all, the following two conceptual distinctions are emphasised to foster the understanding of theoretical and methodological lines of reasoning in this research. Firstly, both 'consumer' and 'customer' are used in literature to describe the buyer and user of a product. As this study focuses on preference measurement of product end-users, they are referred to as consumers rather than customers, following the definition of Blythe (2008, p.5): *"customers are the people who buy the products; consumers are those who consume it"*. The second important distinction in this research concerns the concepts of product 'features' and product 'attributes'. Whereas the terms are used interchangeably in previous research to describe product characteristics, this study distinguishes both concepts to refer to different dimensions of product characteristics. As indicated in the introduction, this study defines *product features* as the objective, manufacturer-provided product characteristics, and *product attributes* as the subjective, consumer-provided product characteristics (e.g., Thompson et al., 2005; Abraham-Murali & Littrell, 1995). The increase of product features thus refers to (engineering) characteristics added by product developers. The attribute categories capability, usability and maintainability thus indicate consumers' perceived product characteristics, expressed and evaluated in online reviews. In the remainder of this chapter, the concepts are further elaborated.

2.1 Online product reviews

Online product reviews are consumers' product evaluations that are freely accessible through the internet. This study focuses on reviews that are posted on designated product pages, rather than on social media, blogs or vlogs. Those reviews can be found on the manufacturing firm's webpage (e.g., Philips.nl), retailers' web shops selling the product (e.g., Bol.com, Coolblue.nl) as well as independent product comparison websites (e.g., Kieskeurig.nl). Figure 2.1 on the next page shows an example of a product review of one of Philips FAEMs, which was posted on a retailer's webpage. Reviews usually consist of five parts that together comprise a consumers' opinion about the product in use: (1) overall product rating, (2) review title, (3) name of the reviewer and posting date, (4) pro and con sections, reflecting the perceived benefits and costs, and (5) free text area, describing the consumer's experiences with the product. Some review pages also include demographic data of the reviewer, such as age, sex, and location. Furthermore, an option to recommend or not recommend the product is featured in some review formats.

Online reviews have become a valuable information source for consumers as well as firms regarding product quality. Research from 2009 shows that at that time, already 70% of consumers retrieved online product reviews during their purchase decision process (Ante, 2009; De Maeyer, 2012). Accordingly,

previous research demonstrated that reviews are a good predictor of word-of-mouth (Zhu & Zhang, 2010), which strongly impacts consumer decision making (e.g. Huang & Chen, 2006; Hardey, 2011) and consumer buying behaviour (e.g. Cheung & Thadani, 2012; Filieri, 2015). Other studies found that firms can benefit from online reviews for predicting product sales (Chevalier & Mayzlin, 2006; Zhu & Zhang, 2010), for product pricing (Chen et al., 2011), competitor analysis (Li et al., 2011; Xu et al., 2016), market structure analysis (Netzer et al., 2012) and marketing strategy (Clemons et al., 2006; Chen & Xie, 2008). Online product reviews have been shown to be highly predictive of consumer needs as well, which can be analysed to arrive at important insights for NPD (e.g., Netzer et al., 2008; Decker & Trusov, 2010). This theoretical background continues focusing specifically on those online review implications that are tailored to improve NPD.

Figure 2.1: Example of a Dutch online product review (source: Coolblue, 2017)

2.2 Online product review analysis for NPD

This subsection discusses online product review research that aims to improve new products. Firstly, advantages of online product review analysis over traditional consumer research methods for NPD are discussed. Secondly, different steps of pre-processing online review data used in previous research are described. These steps include attribute identification, opinion extraction and sentiment analysis. Thirdly, several applications used to improve NPD are discussed, as well as the unique application of online review data in this study.

Analysing online product reviews to improve new products can replace traditional consumer research methods, providing multiple advantages. As opposed to traditional survey-based self-explicated methods, conjoint analysis, and focus group approaches, consumer reviews are based on actual purchase and usage situations and are submitted voluntarily. Consumer needs embedded in reviews are therefore considered more authentic, reliable and representative of the VOC than those collected through surveys or conjoint analyses (Bickart & Schindler, 2001; Ghose & Ipeirotis, 2007). Furthermore, these traditional methods require many resources in terms of time and money (Xiao et al., 2016), which are often limited in NPD processes (Khurana & Rosenthal, 1998; Cooper et al., 2004). In contrast, product reviews are freely accessible on the internet and can be collected and analysed within a relatively short period of time (Xiao et al., 2016). Online product review analysis could therefore provide a solution to many firms to optimise new products, and increase consumer satisfaction and product success.

As can be learned from previous online product review research, the development of appropriate analysis tools and methods is crucial since great amounts of data must be transformed into comprehensible, useful strategic insights (Lee & Bradlow, 2011). The end-objective for product developers is to link these consumer needs to specific engineering characteristics to improve the new product (Qi et al., 2016). Although specific applications for NPD vary across online review studies, many share similar data pre-processing steps (e.g., Decker & Trusov, 2010; Kangale et al., 2016; Qi et al., 2016; Wang & Wang, 2014). These pre-processing steps are shown in Figure 2.2 and described below: attribute identification, opinion extraction and sentiment analysis. Furthermore, different applications for NPD are described.

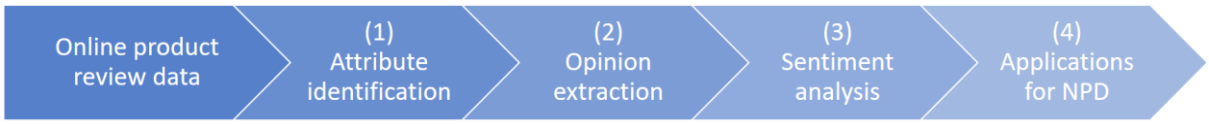

Figure 2.2: Pre-processing steps (1-3) and analysis applications (4) in online review research

The first indispensable pre-processing step in online review analysis is attribute identification. Identifying attributes enables learning specific, actual consumer needs regarding those attributes in subsequent steps of the analysis (Decker & Trusov, 2010). To ensure the integration of the VOC in NPD, product attributes should be identified from consumer's product evaluations in online reviews, rather than using pre-defined manufacturer-provided product feature lists (Jin et al., 2014). A product should thus be assessed as a bundle of consumer-defined attributes, each having its distinct sentiment, and each contributing differently to overall satisfaction (Kotler & Keller, 2006). In online review research, topics that are discussed most often by reviewers are regarded to as most important product attributes (Archak et al., 2011). Several previous studies apply natural language processing (NLP) techniques to content analysis of online product reviews. A popular, semi-automatic method for identifying candidate attributes from online reviews is part-of-speech (POS) tagging. POS tagging is a natural language processing (NLP) technique that labels the grammatical function to each word (Collobert & Weston, 2008). Product attributes often appear as nouns or noun-phrases in product reviews and can therefore easily be extracted from the POS tagging results (Kangale et al., 2016). However, POS tagging for less common languages requires expensive tools, extensive manual guidance or model training (Garrette et al., 2013; Dash, 2013). Another option is to manually identify attributes from the online review text or pro and con sections (e.g., Decker & Trusov, 2010; Qi et al., 2016). Although an appropriate POS tagger is highly time-efficient compared to manual attribute identification, POS tagging cannot reach accuracy and validity levels of manual attribute identification (Hu & Liu, 2004[1]). When the review format used includes pro/con sections, attribute identification, either (semi-)automatic or manual, is often based upon the phrases listed in these pro/con sections rather than on free text areas. Pro/con sections were demonstrated to provide a summary of the free text area (Branavan et al., 2009; Kim & Hovy, 2006). Therefore, attribute identification based on pro/con sections would not cause a significant loss of information and is more efficient compared to free text areas (Xiao et al., 2016; Decker & Trusov, 2010).

The second pre-processing step is to extract consumers' opinions about those attributes that have been identified in the first step. Consumers often value distinct product attributes differently, causing them to express different opinions towards attributes in their product evaluations (Liu et al., 2005). It is important to capture the difference in opinions from online review analysis, since these reflect product strengths and weaknesses that should be accounted for in NPD (Wang & Wang, 2014). Multiple methods can be distinguished for extracting consumers' opinions about attributes from online reviews. Most frequently

applied methods are based on the pro and con sections, and again, NLP techniques. Using pro and con sections for opinion extraction includes simply reporting whether an attribute is mentioned in the pro or con section of the review (Jin et al., 2016[1,2]; Lee & Bradlow, 2011; Decker & Trusov, 2010; Xiao et al., 2016). (Semi-)automatic NLP techniques, most frequently a POS tagger, are applied to extract consumers' opinions from free text areas (e.g., Kangale et al., 2016; Li et al., 2014, Ullah et al., 2016; Wang & Wang, 2014). Opinion words generally appear as adjectives near to the product attribute they refer to. Therefore, adjectives near to nouns (identified as attributes) are collected and paired with the attribute. Again, when pro/con sections are available, they provide the most efficient way to extract consumer opinions from reviews (Decker & Trusov, 2010; Xiao et al., 2016). However, studies that aim to design more elaborate, tailored product improvement strategies seem to choose free text areas over pro/con sections for opinion extraction. If pro/con sections are a summary of the free text area (Branavan et al., 2009; Kim & Hovy, 2006), it can vice versa be argued that free text areas include more elaborate descriptions of consumer needs, aligning with more detailed product improvement objectives.

The third pre-processing step is sentiment analysis, i.e., assigning a sentiment score to each opinion. Sentiment analysis, also referred to as opinion mining, refers to identifying subjective information from data sources, specifically, the attitude of a speaker or writer towards a specific topic (Bruce & Wiebe, 1999; Wang & Wang, 2014). Sentiment analysis is a well-established and widely used method for scoring consumer opinions. In the context of online reviews, sentiment analysis aims to assign scores to consumers' opinions about a product attribute (Wang & Wang, 2014). There are two types of sentiment analysis that are applied in online review research: sentiment polarity (i.e., positive or negative) and sentiment strength analysis (i.e., how positive or negative) (Pang et al., 2002). Sentiment polarity is applied in studies that use pro and con summaries for extracting opinions, indicating whether a consumer has a positive or negative opinion about an attribute (Jin et al., 2016[1,2]; Lee & Bradlow, 2011; Decker & Trusov, 2010; Xiao et al., 2016). For each review, an attribute can either be mentioned in the pro category or con category, resulting in a positive (1) or negative (-1) opinion score (e.g., Decker & Trusov, 2010). This results in a set of both positive and negative evaluated attributes, representing the consumer's evaluation of a product (Das & Chen, 2007). Some studies also add a neutral score for attributes that are not mentioned in a certain review (e.g., Qi et al., 2016). Sentiment strength analysis is often applied in studies that used the free text area for opinion extraction (Qi et al., 2016; Jin et al., 2014; Wang & Wang, 2014; Goswami & Tiwari, 2014). In sentiment strength analysis, each opinion obtains a sentiment score, indicating the extent to which an opinion is positive or negative, e.g., an 'excellent' feature obtains a higher positive score than a 'good' or 'fine' feature. An example of a tool that can be applied to sentiment strength analysis is SentiStrength (e.g., Wang & Wang, 2014). SentiStrength is an algorithm which is specifically developed to perform sentiment analysis on short texts, using several methods to assign both a positive and negative sentiment strength score to a word, phrase or text (Thelwall et al., 2010). These scores can be computed in an overall score on, for instance, a scale from 0 to 10 (Wang & Wang, 2014). The advantage of a sentiment strength over sentiment polarity is that it enables researchers to rank opinions and attributes based on scores, allowing for more detailed analysis (Goswami & Tiwari, 2015; Wang & Wang, 2014). Disadvantages include the need for expensive, well-trained tools for less common languages and opinion subjectivity, as well as the need for analysing free text areas rather than pro and con sections (Li & Li, 2013).

The fourth step in online review research, following up on data pre-processing, include applications for NPD. Once researchers have extracted attributes and opinions from the selected set of reviews, and determined sentiment polarity or strength, they use these data for their analysis to arrive at important insights for NPD. Data analysis approaches range from econometric preference analysis (Decker & Trusov, 2010), multi-linear regression analysis (Goswami & Tiwari, 2015), ordinal classification

analysis based on pairwise algorithms (Jin et al., 2014), to a feature-based review-summary (Kangale et al., 2016), and correspondence analysis for market structure mapping (Lee & Bradlow, 2011).

This study adds another relevant angle to the above list of applications for product improvement. As described earlier, online review research generally refrains from using product features as topics of consumers' evaluations. However, many firms differentiate their products based on additional, unique features (Thompson et al., 2005). It is therefore considered relevant to include number of features as exogenous variable predicting consumer satisfaction, accounting for consumers' evaluations of product attributes that arise from using the feature-rich product. As product features increase, consumers' perceptions of the product's capability, usability and maintainability might change, affecting consumers' overall satisfaction with the product (Thompson et al., 2005). The product attributes, opinions and sentiment scores obtained through pre-processing online reviews can be clustered into attribute categories to represent consumers' perceptions of capability, usability (Wu et al., 2014) and maintainability. Section 2.3 continues with the describing effect of adding product features on consumer satisfaction, and the role of attribute categories capability, usability and maintainability, based on previous literature.

2.3 The effect of adding product features on consumer satisfaction

Consumer satisfaction refers to the overall post-purchase evaluation (Fornell, 1992, p.11) of the perceived discrepancy between prior expectations (or some norm of performance) and the actual performance of the product as perceived after its consumption (Tse & Wilton, 1988, p. 204). In this study, consumer satisfaction is examined in the context of increasing product features. Paragraph 2.3.1 discusses the 'feature fatigue effect' as demonstrated by Thompson et al. (2005), explaining why companies often choose a feature-based differentiation strategy. Furthermore, it is described why consumers might initially prefer and buy feature-rich products due to high perceived capability, but later put more emphasis on the usability of products, leading to decreased satisfaction. Paragraph 2.3.2 introduces the concept of perceived maintainability, which is added as third attribute category influencing consumer satisfaction for maintenance-sensitive products, in addition to perceived capability and usability.

2.3.1 Feature fatigue: perceived capability vs. perceived usability

With the objective to enhance consumer satisfaction, firms in many industries increasingly tend to differentiate products by adding features (Mukherjee & Hoyer, 2001; Goldenberg et al., 2003; Brown & Carpenter, 2000). Feature-based differentiation strategies, as well as the widely-accepted conjoint analysis methods, suggest that each positively valued feature increases consumer satisfaction and product market share (Nowlis & Simonson, 1996; Thompson et al., 2005). An often-overlooked issue is, however, that the cost arising from the complexity to use a feature-rich product often exceeds the benefit of having many functions (Burke, 2013; Mick & Fournier, 1998). Thompson et al. (2005) refer to this phenomenon as feature fatigue, and examined its effects on consumer satisfaction. They find that higher levels of features increase satisfaction through increased *perceived capability*, i.e., consumers' perceptions of the product's ability to perform the desired functions. But at the same time, higher levels of features may decrease satisfaction through decreased *perceived usability*, i.e., consumers' perceptions of the ease of learning and using the desired functions (Thompson et al., 2005). In the absence of the usage perspective, consumers tend to desire and purchase feature-rich products as perceived capability is high. After purchasing a feature-rich product, consumers often appear to find it difficult or stressful to use particular features (Thompson & Norton, 2011; Thompson et al., 2005; Mick & Fournier, 1998; Wiklund, 2012; Nielsen, 1994, p. 155). They may even fail or refrain to use features that they initially

desired (Goodman & Irmak, 2013), leading to decreased satisfaction. Building on these findings, Rust et al. (2006) provide guidelines for firms to defeat feature fatigue. To reduce product complexity, they recommend firms to develop a variety of simpler products, each targeting a particular consumer segment. Additionally, Rust et al. (2006) suggest using tools, recommendation agents and product trials to help consumers make the right purchase decision by forcing them to consider which features they will actually need and use.

Thompson et al. (2005) conceptualised the product's ability to meet consumer needs, influencing consumer satisfaction, as combination of perceived capability and perceived usability of the product. In online review research, the concepts of perceived capability and usability can be regarded to as categories of product attribute evaluations. This study adds perceived maintainability as additional attribute category, since maintenance-sensitive products are the subject of research.

2.3.2 Perceived maintainability

As described in the introduction, for specific product types, another determinant of consumer satisfaction needs to be assessed: maintainability. Maintainability applies to products that require regular maintenance by the user, rather than maintenance as a service by a firm, e.g., printers, washing and drying machines, complex kitchen and coffee machines, such as FAEMs. Consumers may require maintenance operations (e.g., replacing an ink cartridge) to be performed quickly, easily and with no training (Hooks & Farry, 2001, p.76). Accordingly, this study defines *perceived maintainability* as consumers' perceptions of the ease of cleaning and maintaining the product to keep the desired functions in working condition. Especially, since sustainability has become an important focus of government, companies and society, products are increasingly designed for maintainability rather than disposability (Karwowski et al., 2011). As noted in the introduction, no prior research was found to examine the effect of perceived maintainability on consumer satisfaction for physical products. The following sections provide an overview of prior research on maintainability, establishing theoretical foundations for the link between number of features and maintainability as well as maintainability and satisfaction.

Slavila et al. (2005) emphasise the importance of integrating maintainability aspects early in the product development stage for products that require maintenance. They state that, if the required maintenance by the user is difficult to perform, more time and effort must be devoted, increasing product life cycle costs and decreasing availability and reliability of the product. For a product to be highly maintainable, the design should be simple in a way that components are easy to access, remove, and replace, and few maintenance products are required (Pecht, 2009). As both Slavila et al. (2005) and Pecht (2009) link the need for maintainability to the complexity of products, it is assumed that maintainability becomes more of an issue when complexity of a product, i.e., the number of features, increases.

In the research area of software-based products, maintainability by the user has been acknowledged as factor of consumer satisfaction for decades (e.g., Boehm et al., 1976; Kekre et al., 1995). From the perspective of component-based software systems, maintainability is defined as the ability of the software to be modified (ISO/IEC, 1996; Mari, 2003). Although physical products cannot literary be modified, similarities can be drawn from the shared goal of keeping the product in working condition. Additionally, for software as well as physical products, the importance of maintainability to perceived product quality and satisfaction should not be underestimated, since maintainability concerns the whole life cycle of the product (Mari, 2003). In the context of clothing, Seva et al. (2007) demonstrated that products that are easily maintained elicit positive emotional reactions from consumers since they imply a convenient and stress-free lifestyle. These findings indicate that perceived maintainability can be regarded to as factor of consumer satisfaction.

For complex (i.e., feature-rich) physical products that require frequent and extensive maintenance by the user, the effect of perceived maintainability should be assessed in consumer research. NPD may otherwise fail to address specific, latent, needs arising from maintaining the product. Instead, product developers may prioritise capability, usability, or even ease of development (Hooks & Farry, 2001, p.153), harming consumer satisfaction and product success. This study aims to close the gap in literature by integrating perceived maintainability in the model proposed by Thompson et al. (2005). The proposed conceptual model and hypotheses are described in section 2.4.

2.4 Theoretical model and hypotheses

The theoretical model proposed in this paragraph connects the concepts discussed previously in this theoretical background. Consumers' product evaluations in the form of online reviews reflect consumers' perceptions of capability-, usability-, and maintainability-related product attributes as well as their overall satisfaction with the product. When these perceptions of the product are compared between products that differ based on the number of features, the effect of adding product features on consumer satisfaction can be assessed. To emphasise the difference between the theoretical concepts used, the definitions of product features, product attributes, perceived capability, usability and maintainability, and consumer satisfaction are summarised in Table 2.1.

Table 2.1: Definitions of the concepts described and used in this study

Concept	Definition used in this study
Product features	The objective characteristics of a product that are exogenously defined by manufacturers, expressed in product descriptions in (online) shops, i.e., mostly technological opportunities of a product that distinguishes it from competitors (Kotler & Armstrong, 2007; Archak et al., 2011; Thompson et al., 2005; Netzer et al., 2008).
Product attributes	The subjective characteristics of a product that are endogenously perceived and evaluated by consumers, expressed in their product evaluations arising from usage situations, i.e., costs and benefits of a product to an individual consumer (Abraham-Murali & Littrell, 1995; Thompson et al., 2005; Kangale et al., 2016). In this study, product attributes are grouped into overarching attribute categories perceived capability, usability and maintainability.
Perceived capability	Consumers' perceptions of the product's ability to perform the desired functions (Thompson et al., 2005).
Perceived usability	Consumers' perceptions of the ease of learning and using the desired functions (Thompson et al., 2005).
Perceived maintainability	Consumers' perceptions of the ease of cleaning and maintaining the product to keep the desired functions in working condition (altered based on Hooks & Farry (2001) and Seva et al. (2007) to fit with the definitions of capability and usability by Thompson et al. (2005)).
Consumer satisfaction	To the overall post-purchase evaluation (Fornell, 1992, p.11) of the perceived discrepancy between prior expectations (or some norm of performance) and the actual performance of the product as perceived after its consumption (Tse & Wilton, 1988, p. 204).

Figure 2.2 presents the conceptual model that was developed. The model is designed to measure the effect of adding product features on consumer satisfaction, through consumers' perceptions of capability, usability and maintainability, and extends the model of Thompson et al. (2005), adding maintainability as third mediator. Based on the significant findings of Thompson et al. (2005) for the

mediating effects of capability and usability in their model, this study assumes the mediating effects for these two variables to be similar in terms of direction. Thompson et al. (2005) found that, after consumers had used the product, number of features had a positive effect on capability, but negative effect on usability. Both capability and usability have a positive effect on consumer satisfaction. Since the effect of usability on satisfaction was higher than that of capability, the overall indirect effect of number of features on consumer satisfaction was negative, indicating feature fatigue. It should however be noted that the type of product researched in this study, the FAEM, is different from those studied by Thompson et al. (2005): digital audio and video players. Furthermore, Philips has been focusing increasingly on creating products that are not only advanced in terms of technology, but also customer-focused and user-friendly, also implied by their former slogan 'Sense and Simplicity' (NOS, 2013). These circumstances might lead to divergent results of the role of capability and usability in the features-satisfaction relationship.

Within the features-satisfaction relationship in the conceptual model, the mediating role of maintainability is assumed to be similar to the role of usability in the study of Thompson et al. (2005). Like usability, maintainability is difficult to assess when buying a product. After purchasing and using a feature-rich product, maintainability is likely to become increasingly important to consumers as maintenance will be required to keep the product in working condition. Due to the increased complexity of a product inherent to adding product features (Rust et al., 2006), consumers are more likely to perceive maintenance of the product as overwhelming and difficult (Slavila et al., 2005; Pecht, 2009). These theoretical findings comprise Hypothesis 1.

H₁ *As the number of features included in a product increase, perceptions of the product's maintainability decrease.*

As described earlier, maintainability by the user has been considered an important determinant of consumer satisfaction in the research area of software-based products for decades (e.g., Boehm et al., 1976; Kekre et al., 1995). When products are easily maintained, consumers experience positive emotions towards the product (Seva et al., 2007), causing them to be more satisfied with the product. Therefore, like capability and usability (Thompson et al., 2005), maintainability is perceived to influence consumer satisfaction in a way that higher maintainability leads to higher satisfaction. These theoretical findings provide the basis for Hypothesis 2. The hypothesised effects are also indicated in the conceptual model in Figure 2.2.

H₂ *As perceived maintainability increases, consumer satisfaction increases.*

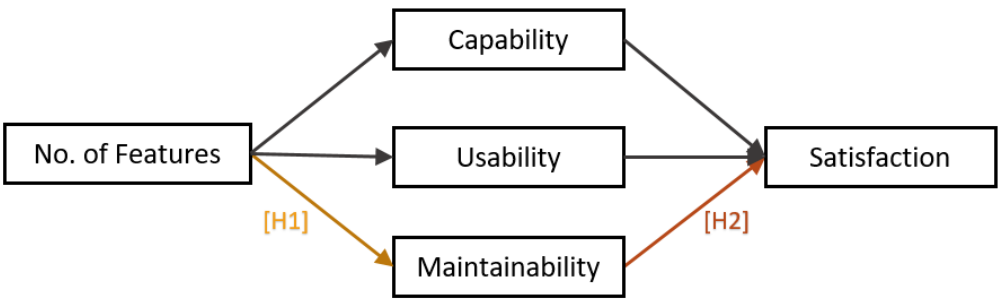

Figure 2.2: Conceptual model of the study, extending the model of Thompson et al. (2005) with maintainability as additional mediator

3. Methodology

This chapter builds forward on the theoretical background, describing how the effects within the conceptual model are operationalised and measured. Methodological issues are clarified and decisions are motivated. Section 3.1 discusses the research design used, describing the extensions of the research by Thompson et al. (2005), presenting a research framework outlining all methodological steps, the case study used and the research settings. Section 3.2 describes the data collection process, section 3.3 the data pre-processing steps, and section 3.4 the data analysis that was applied to arrive at important insights to improve new maintenance-sensitive products.

3.1 Research design

The research design is based on the article by Thompson et al. (2005): "*Feature fatigue: when product capabilities become too much of a good thing".* They examine the role of increasing product features on consumer satisfaction, accounting for the mediating role of perceived capability and usability. This study adopts the method of Thompson et al. (2005), extending it in three ways: by using online review data rather than survey data to measure consumer needs, by adding perceived maintainability as third - hypothesised - mediator, and by measuring the effects of distinct product attributes on satisfaction. These extensions are more elaborately described in the remainder of this methodology chapter. The research framework in Figure 3.1 provides an overview of the various methodological steps conducted, which are discussed in sections 3.2-3.4. During the data pre-processing phase (3.3), three different validation sessions were conducted, as indicated in the framework. Different users of Philips' FAEMs were approached and asked to identify, categorise and score product attributes. This use of multiple sources of evidence is called triangulation. Triangulation ensures converging lines of inquiry, verifying repeatability of the steps conducted (Stake, 2000), which increases reliability and validity of the research (Yin, 2013). Reliability refers to the consistency of measures, and validity refers to the extent to which a measure of a variable really measures that variable (Bryman, 2015; p.168-174).

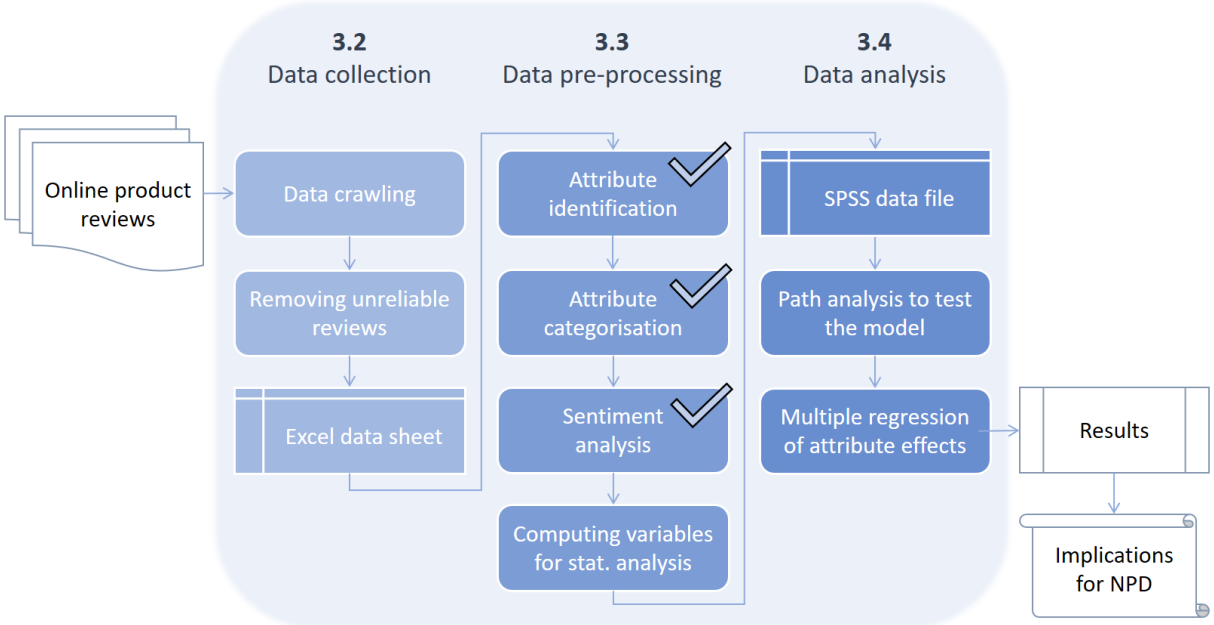

Figure 3.1: Research framework

The case study that is used to conduct the steps in the research framework is that of Philips's FAEMs. A single-case study approach was chosen, since it allows for studying the relationships between number of features, capability, usability, maintainability and satisfaction within a real-life context (Yin, 1981). The research question was addressed based on only one case, since each product category has unique characteristics and each firm's NPD is unique in its combination of strategy and environment (Porter, 1998). This will ensure a clear way of analytic generalisation (Yin, 2013). The case study concerns Philips' FAEMs that are currently available on the Dutch coffee machine market. FAEMs have shown initial growth in Western European coffee machine market over the past years and have the potential to capture a great market share in the near future (Morris, 2013; Businesswire, 2016). This case is selected as FAEMs are generally differentiated based on added features arising from advanced technologies, and both firms and consumers increasingly adhere to maintainability, on top of capability and usability (F. Meinster, personal communication, March 6, 2017). FAEMs are increasingly featured with advanced operating, grinding, brewing, and milk frothing technologies. Compared to drip filter and single-serve coffee machines, extensive and frequent maintenance is required by the user to keep a FAEM working properly (Philips, 2017). Philips, recognises consumers' needs for maintainability. The problem that Philips faces, however, is that it lacks knowledge of consumers' perceptions of capability, usability and maintainability for different levels of product features and their relative effects on consumer satisfaction (F. Meinster, personal communication, March 6, 2017). This knowledge is considered crucial to Philips' objective to once again obtain a dominant design position (Christensen et al., 1998) in the coffee machine market, as did their Senseo in the 2000s (Griffioen, 2008; Van Kralingen, 2007). Each added feature in fact provides an additional opportunity to satisfy consumers through perceived capability. However, when consumers start using the feature-rich product, they might be overwhelmed by the number of options to choose from, finding it difficult to learn and use all features (i.e., decreased usability). Furthermore, when the number of features increases, more extensive and more frequent maintenance might be required which can be perceived by consumers as difficult or annoying (i.e., decreased maintainability). Hence, adding product features might in fact decrease satisfaction through usability and maintainability. Whether the feature fatigue effect exists for Philips' FAEMs is assessed in this study, as are the relative effects of perceived capability, usability and maintainability. Recommendations focus on how FAEMs should be improved to maximise consumer satisfaction and, in turn, the product's success on the market.

It was decided to perform this study based on three product types within Philips' FAEM product line to create a research setting similar to that of Thompson et al. (2005), allowing for appropriate comparison of the results. While the products studied by Thompson et al. (2005) included 7, 14, and 21 features, an equal distance between the number of features of Philips' FAEMs could not be attained. This is due to the low number of online product reviews available for mediate FAEM types. To still represent an increase of product features, online product reviews of these three types of Philips FAEMs were collected: the HD8821, HD8824 and HD8834, including 14, 16, and 21 features, respectively. The product types and features are listed in Appendix A.

3.2 Data collection

In contrast to the experimental settings and surveys applied by Thompson et al. (2005), this study is based on real usage situations of actual consumers as reflected by online product reviews. As discussed earlier, online reviews offer a promising new information source, allowing for more reliable, effective and efficient consumer research for integrating the voice of the consumer in NPD, fostering successful product innovation (Bickart & Schindler, 2001; Ghose & Ipeirotis, 2007; Archak et al., 2011). The data collection phase consists of two steps, which are discussed below. Online reviews were first crawled

(i.e., copied) from product pages into an Excel file, followed by the removal of unreliable reviews to ensure reliability of the data.

The first step of data collection was crawling online product reviews from webpages. Reviews of the three selected Philips' FAEMs were retrieved from the Philips product page (Philips.nl), Dutch retailer websites (Bol.com, Coolblue.nl, Mediamarkt.nl) as well as one comparison website (Kieskeurig.nl). Although each website uses a slightly different review format, they are all featured with reviewer name, review title, pro section, con section, free text area, and overall product rating. Some websites also report age, sex and location of the reviewer, the time of owning the product, and whether the reviewer would recommend the product to others. All available review data was manually crawled (i.e., copied) from the websites and collected in an Excel sheet. The reason for choosing manual scraping over automatic scraping tools is that automatic tools are not free of charge and not suited to every webpage (Abrahams et al, 2013; Wang & Wang, 2016; Qi et al., 2016). When a firm like Philips would decide to conduct online review analysis on a regular basis, it is recommended to try different automatic data crawling tools or have one developed for each targeted webpage. A screenshot of the resulting Excel datasheet is shown in Figure 3.2. The columns represent the different parts of the review, and each row represents one unique review. As can be seen from this sheet, each review was assigned a unique review number to be able to retrieve particular reviews in a later stage. Furthermore, the product type was listed for each review. In total, 623 reviews were collected, which is also shown in Table 3.1.

	A	B	D	E	J	K	L	M	N	O
1	Review no.	Product type	Source	Username	Posting date	Satisfaction (star rating)	Review title	Free text area	Pros	Cons
236	235	HD8824/01	Coolblue	Dick Dijk	1-12-2016	4	prijs kwaliteit ruimschoots in verhouding	bevalt goed. mooie droge sjoelbakschijven waardoor veel minder vervuiling.	weinig lawaai bedienings gemak smaakinstelling	water bijvullen ivm bovenkastje
237	236	HD8824/01	Coolblue	Dennis van Marle	14-11-2016	4	Mooi en goed product	Produkt bevalt. prijs kwalitiet gewon top t.o.v. de minpuntjes.	mooi smal toch wel	melkslang kan niet blijven hangen bij reiniging sterkte moeilijk te wijzigen
238	237	HD8824/01	Coolblue	Carolien Janssen	9-10-2016	5	Mooi en makkelijk in gebruik!	Het duurde twee weken voordat de koffie op smaak was. maar nu is de koffie perfect. Erg makkelijk in gebruik door duidelijke visuele weergaven. Wij zijn er erg blij mee.	Makkelijk in gebruik Lekkere koffie Duidelijke weergave wanneer iets op is of gereinigd moet worden	
239	238	HD8824/01	Coolblue	martijn vossers	4-9-2016	5	Goed en simpel.	Simpel overzichtelijk apparaat. makkelijk te reinigen. Eerst dacht ik dat de koffie niet erg warm was. maar dat heb ik gemeten en is altijd iets boven de 80 graden. gewoon goed dus. Snelle opwarmtijd. Automatische stand-by regeling	Makkelijk te reinigen Simpele overzichtelijke bediening Geen overbodige instellingen, gewoon goed.	Sommige kunstof delen slap uitgevoerd, zoals draagbeugel waterreservoir, maar makkelijk op lossen door water bik te vullen met een maatbeker bijvoorbeeld.

Figure 3.2: Screenshot of the collected data in Excel

The second step of the data collection included the removal of unreliable reviews from the database to ensure reliability of the data. The detection of spam and fake reviews has received considerable attention in previous online review research (e.g., Kangale et al., 2016; Zhou & Duan, 2015). As all product selling pages used in this study (Philips.nl, Bol.com, Coolblue.nl, Mediamarkt.nl) only allow actual purchasers of products to submit reviews from their user accounts, no spam and fake reviewers need to be detected. Furthermore, independent comparison page Kieskeurig.nl guarantees high quality and

reliability of reviews as each review is manually checked and purchase receipts are requested on a sample basis (Kieskeurig, 2018). It was assumed that Kieskeurig's own spam filter only allows trustworthy reviews. Advanced spam detection methods were therefore not needed in this study. However, to further refine the data and increase reliability, three different, additional filtering methods were applied. These methods and reasons for removing reviews are described below: removing incentivised reviews that are based on simulated usage situations, removing Belgian reviews that were copied to the Dutch manufacturer page, and removing (almost) identical reviews that occurred on more than one review page.

Firstly, incentivised reviews of simulated usage situations were excluded, based on the findings of prior research (Mayzlin, 2006; Mayzlin et al., 2014; Petrescu et al., 2007). Firms often carry out marketing campaigns in which panel members are asked to write a review in turn for a free product or trial usage (Mayzlin, 2006). Incentivised reviewers tend to give higher overall evaluation scores than non-incentivised reviewers, leading to deceiving outcomes (Reimer & Benkenstein, 2016; Petrescu et al., 2017). Furthermore, those incentivised reviewers are not actual buyers and users of a product and therefore unable to represent real consumer preferences (Srinivasan et al., 1997). Amazon has recently started to remove incentivised reviews for the above reasons (Perez, 2016). Accordingly, to ensure reliability of the data, it is decided to exclude incentivised reviews by simulated consumers from the dataset. Philips regularly organises campaigns to generate more (positive) online WOM. On the Philips webpage, these incentivised reviews are marked with 'reviewer participated in a Philips product test'. These panel members did not purchase the product and were rewarded to use the product and write a review. Therefore, these incentivised reviews do not reflect a real but rather simulated purchase and usage situation. In total, 17 reviews were removed. For a review system to be effective, consumers must be motivated either by the environment, or by incentives within the system (Hu et al., 2006). Many retailers hence encourage their consumers to write a review, for instance, by offering prizes to be won or small discounts on the next purchase (e.g., Bol.com, Coolblue.nl). Although these reviewers are also incentivised to some extent, they are in fact verified consumers who have actually purchased and used the product on their own behalf (Bol.com, 2018; Coolblue, 2018). Therefore, these reviews are still perceived as highly trustworthy and therefore not excluded from the dataset.

Secondly, other reviews from Philips.nl that were excluded from the analysis, are those marked with 'written by a consumer visiting Philips.be'. This implies that the review was submitted on the Belgian version of the Philips website and was copied to the Dutch website. The decision to exclude those reviews is made because this study focuses on FAE users in the Netherlands only, as geographical and cultural differences between coffee preferences exist (Morris, 2013) and Belgian reviews might therefore bias the data towards divergent opinions. In total, 15 reviews were excluded.

Thirdly, (almost) identical reviews that occurred on more than one retailer page were removed, based on awareness raised by previous researchers (Zheng et al., 2013; Yan et al., 2015). Duplicate and near-duplicate reviews are considered as spam, contaminating the data, and therefore need to be detected and removed (Jindal & Liu, 2008). At comparison website Kieskeurig.nl some reviews were marked with 'originally posted on Philips', implying that the review is just a copy of one that is already included in the database, as Philips.nl reviews were collected first. Only one of these review duplicates was included in the database. Furthermore, some reviews or reviewer names seemed familiar during the scraping process. Therefore, a check for duplicate and similar reviews was performed in Excel. 13 reviews were removed accordingly, as they were (almost) literally copied to different review websites, leaving only unique consumer reviews in the database. When a single consumer's product evaluation was posted on and collected from multiple review pages, the perceived capability, usability, maintainability and satisfaction scores would also be measured multiple times. When computing the overall variable scores

for the product type in question, double weight will be given to the opinion of that single consumer, compared to other consumers' opinions. Hence, the data would be biased towards that consumer's opinion and therefore less representative of aggregated evaluations of the consumer population. It is therefore not desirable to include multiple reviews from the same writer.

Table 3.1 provides an overview of the numbers of reviews collected, removed in each exclusion step, and final numbers of usable reviews for each product type. While in this current study, no big amount of data was removed in the exclusion steps, it was still considered important to describe these steps accurately. This ensures repeatability of the steps as well as the applicability to different, possibly bigger data sources, enhancing reliability and validity of the approach. In future research based upon different data, researchers might be confronted with bigger amounts of unreliable review data. Based on the steps discussed here, these data can be excluded from the dataset in a structured way.

Table 3.1: The number of reviews collected, removed and used for further analysis for each product type

Product type	Features	Reviews collected	Reviews removed			Reviews used
			Incentivised	Belgian	Duplicated	
HD8821	14	351	12	11	9	307
HD8824	16	180	5	0	4	170
HD8824	21	92	0	4	0	88
Total		623	17	15	13	565

From the final review data in the Excel sheet, review characteristics could be obtained, as well as demographics of the reviewer population, which are listed in Appendix B. Review characteristics include the average number of pros (2.78) and cons (1.18) that reviewers mention, and the mean satisfaction score (3.96). Furthermore, the percentage of reviewers who would recommend their FAEM (84.89%) who would not (15.11%) is reported, of which data was however not available on each review page. For the demographics of the reviewer population, it should be noted that these data were only available for 5.66-18.58% of the reviews collected, due to the different review formats used on the webpages. Based upon the available demographics data, the average age of the reviewer population is 44.85, 66.67% are male; 33.33% female, average time the product is in use is 1.38 months. Characteristics and demographics on product-level for each source are also computed and listed in Appendix B.

3.3 Data pre-processing

To enable measurement of consumer needs from online reviews, the data required pre-processing (Xiao et al., 2016). The textual review data that was collected was converted into numerical data that allows for quantitative analyses of the relationships in the conceptual model. Based on previous online review research (Decker & Trusov, 2010; Xiao et al., 2016; Lee & Bradlow, 2011; Jin et al., 2016[1,2]), appropriate pre-processing methods were selected, which are described below: identifying product attributes from pro and con sections (3.3.1), categorising the identified attributes into the categories capability, usability and maintainability (3.3.2), scoring opinions that consumers have towards product attributes using sentiment polarity analysis based on pro and con sections (3.3.3), and finally, computing the variables in the model based on the data for statistical analysis (3.3.4).

3.3.1 Attribute identification

The goal of attribute identification is to find the product attributes that consumers evaluate in their review, i.e., perceive as important (Qi et al., 2016). It was decided to use pros and cons sections for attribute identification as it provides a simple, comprehensible, but effective approach (Jin et al., 2016). Using free text areas would imply a learning approach, requiring a great amount of labelled data for model training, which is time-consuming as well as out of the scope of this research. As the pro and con section generally provides a summary of the free text area (Branavan et al., 2009; Kim & Hovy, 2006), using pros and cons as a source for attribute identification would not cause a significant loss of information compared to using free-text areas (Xiao et al., 2016). Furthermore, pro and con lists appear in structured and semi-structured review formats. Semi-structured review formats are widely used by many online product review platforms and manufacturers' websites (Xiao et al., 2016). This implies that attribute identification based on pros and cons phrases can be applied to online reviews collected from various platforms, enhancing generalisation of the research approach.

It was decided to manually identify product attributes from the pro and con sections. Although semi-automatic POS tagging is a common, appropriate approach for extracting product attributes from pro/con sections (e.g., Kangale et al., 2016; Li et al., 2011), trial and error found that manual attribute identification better suited the data. Available POS tagging tools are either expensive or not well suited to the Dutch language. Freely available TreeTagger was used to try and obtain POS tags from the pro and con phrases. Since this tool features a Dutch parameter set, it was initially expected to deliver a list of frequently mentioned product attributes (i.e., nouns) easily and quickly. However, the tagging mechanism was unable to detect product attributes consisting of multiple words. For instance, "*the sound of grinding the coffee"* was tagged as two separate nouns, *"sound"* and *"coffee"*, and one verb, *"grinding*. As this is one of many examples, it would still require tremendous manual processing or model training to detect, combine and group multiple-word attributes. Hence, manually identifying product attributes from pro and con phrases was chosen as it is free, less complex and more effective.

Manual identification of attributes consists of reading the pro and con phrases, and adding an attribute name for each conceptually distinct product attribute mentioned to the dataset. This attribute identification procedure was adapted from Decker & Trusov (2010), Qi et al. (2016), and Archak et al. (2011). Examples of pro and con phrases and labelled attribute names are shown in Table 3.2. Similar attribute names were grouped, and assigned a shared attribute to reflect all included attributes. The following step was added to the attribute identification approach, extending previous methods. When the attribute mentioned in the pro/con section was not clearly defined, titles and free text areas of the review were consulted to retrieve more detailed attribute descriptions. In that way, many attributes could still be defined and analysed, increasing accuracy of the approach. In total, 84 product attributes were identified from the pro/con sections, which are listed in the right part of Appendix C.

Table 3.2: Examples of pro and con phrases labelled with attribute names

Pro/con phrase	**Attribute name(s)**
Pro: *"delicious coffee within one minute"*	Coffee taste/quality; Speed of making coffee
Con: *"requires maintenance frequently"*	Frequency of maintenance
Pro: *"beautiful design"*	Design (aesthetics)
Con: *"cleaning after using the milk frother"*	Milk frother cleaning
Pro: *"operating the machine is convenient"*	Ease of operation
Con: *"the quality of the coffee is bad and temperature too low"*	Coffee taste/quality; Coffee temperature

Because a great deal of subjective judgement is involved in attribute identification, reliability of this step was assessed based on inter-observer consistency (Bryman, 2012, p.170). The attribute identification step was therefore validated by a user of a Philips FAEM who was found in the researcher's personal network. This Dutch, 26-year old, highly-educated male is a frequent coffee and espresso drinker, who has been using his Philips HD8821 FAEM at least twice a day for the last seven months. Based on his profile and expertise with the product group in question, this user was considered a highly-reliable source for validation.

Validation was performed based on a subset containing 85 out of 565 randomly selected reviews (~15%) from the dataset, considering the limited time and resources available to this study as well as the availability of the validator. Based on the author's experience, processing the entire data set takes approximately 20 hours to complete. 85 reviews would still require approximately 3 hours, which was considered manageable. It is however important to guarantee reliability of this step. Hence, the expected percentage of total reviews that could be identified from 85 reviews was calculated. The formula of Nielsen & Landauer (1993) provides a widely-applied, suitable method to determine the number of users needed in user tests (Sauro & Lewis, 2016). In the context of this study, the number of reviews represent the number of users that 'tested', i.e., used, the product, and the number of issues detected is the number of pros and cons mentioned in the review. It was established that 98.35% (82.61 attributes) of all identified product attributes from the complete dataset (84) is to be covered by 85 consumer reviews. This calculation is based on the average number of product attributes addressed by a single consumer (3.96). The complete calculation is included in Appendix D.

The validator was asked to transform the pro/con phrases of those 85 reviews into product attribute names. Detailed instructions for the validation task, as were sent to the validator, are included in Appendix E. In cooperation with the validator, the attribute list resulting from the validation task was compared to the initial set of product attributes as identified by the researcher. Three separate attribute lists were computed, as displayed in Appendix C: researcher's attributes matched to the validator's attributes, researcher's attributes that were labelled as 'external' by the validator, and researcher's attributes that did not occur in the subset according to the validator. The first attribute list computed and discussed was the matched attribute list. The validator generally used a wider scope for defining attributes, i.e., 45 attributes were used to capture the same topics for which the researcher used 74 attributes. For instance, he used 'desired output' to refer to the quality and taste of the full range of drinks that the FAEM in question can make. In one case, the validator identified two different product attributes which were clustered in only one attribute by the researcher: he perceived the 'stand-by mode / on/off-button (light)' as two separate attributes. In his experience, these are in fact two different issues: the (heavily blinking) on/off button during stand-by modus, and the (quite short) time after which the machine automatically switches to stand-by modus. To capture most specified consumer needs from the online reviews in subsequent analysis steps, it was decided to remain most detailed, elaborate attribute labels. This resulted in 75 unique product attributes. Not all attributes identified by the researcher could be matched to the validator's attributes. Out of the 84 attributes originally identified by the researcher, 10 attributes could not be matched to the validator's attributes. These 10 attributes were also evaluated with the validator and divided into the two distinct attribute lists: external and not-occurred attributes. The second list of attributes included 3 attributes that were perceived as external factors by the validator: 'beans (choice/freshness)', 'service (Philips, spare parts, reparation)', 'service (delivery/retailer)'. Although these factors might influence consumer's satisfaction with the overall product experience, they are not directly related by the product which was developed and manufactured by Philips. These external attributes were excluded, since they fall out of the scope of this research, which is to identify improvements of the actual, physical product. The third, remaining, attribute list consists of 7 product attributes identified by the researcher, which were not recognised by the validator, i.e., they did not

appear in the subset of reviews examined by the validator. The validator agreed that these were conceptually different product attributes, reflecting different consumer needs that might impact consumer satisfaction with the product. Hence, these attributes were included for further analysis.

The final list consists of 82 product attributes; 75 researcher-validator matched attributes and 7 additional attributes that did not occur in the validation subset. The proportion of agreement in this validation step is 89.29%, which was computed by dividing the number of matched attributes (75) by total the number of originally defined attributes by the researcher (84). Inter-observer consistency (Bryman, 2015; p.170) is therefore found sufficient, which ensures reliability of the attribute identification method used.

3.3.2 Attribute categorisation into capability, usability and maintainability

Once attributes had been identified, the next pre-processing step was to group these attributes into attribute categories, representing the underlying consumer needs that are studied: capability, usability, and maintainability. This implies that perceived capability, usability and maintainability were measured based on consumers' evaluations of product attributes within these categories. As indicated earlier, categorisation of product attributes is an addition to previous online review research methods for NPD. Attribute categorisation reflects the tendency of consumers to structure their experiences with specific products into categories (Gutman, 1982; Yang, 2013), and thus enables interpretation of consumers' cognitive processes during product evaluation. This categorisation will enable product developers to focus on these underlying consumer needs.

Each previously identified attribute was manually categorised by the researcher as well as five users into either perceived capability, usability or maintainability. Five users of FAEMs were involved for cross-validation as this pre-processing step is crucial to this study and its implications. In this way, inter-observer consistency and face validity were assessed. Inter-observer consistency is explained in section 3.3.1. When a researcher develops a new measure, face validity should be established, implying that the measure actually reflects the content of the concept in question (Bryman, 2015, p.171). Face validity of capability, usability, and maintainability can be established by asking users to determine whether the measure seems to reflect the concept in question (Bryman, 2015, p.171). The five FAEM users were found in the researcher's personal network through social media postings (i.e., WhatsApp and Facebook groups). They are aged between 21 and 57, four male and one female, mediate- to highly-educated, all live in the region of Eindhoven and are frequent users of their FAEMs. Google Forms was used to collect their perceptions of how the product attributes should be categorised. After entering demographics, e-mail address and model number of the Philips FAEM they own, respondents could indicate their categorisation choices in a multiple-choice matrix. In this matrix, product attributes were displayed in arbitrary order. For each attribute, only one answer could be selected out of: "capability", "usability", "maintainability" and "none/multiple categories". More detailed instructions for filling in the Google Form can be read in Appendix F, which were also provided to the respondents of the survey.

The categorisation of an attribute was accepted when at least three respondents had selected the same category. Out of the set of 82 attributes, 72 attributes could be clustered due to agreement between the respondents. Inter-observer consistency and face validity of the attribute categories are therefore assessed as high. However, there were 4 product attributes that respondents did not agree upon: 'accessibility of parts', 'cable (length)', 'drip tray capacity (frequency of emptying)', 'size of the machine'. Either two categories scored 2 out of 5 votes, or the option "none/multiple categories" was selected by multiple respondents. Those 4 attributes were discussed in short telephonic consults with the respondents. Consensus was reached for each product attribute, which could then be categorised as well. Appendix G includes the answers provided in the survey, as well as the final categorisation.

3.3.3 Sentiment analysis

When attributes have been identified and categorised, the next pre-processing step includes extracting the opinions that consumers have towards those attributes. A widely used mechanism for processing subjective texts is sentiment analysis, which was emphasised in more detail in the theory section. Sentiment polarity was chosen over sentiment strength, as sentiment strength would require a tailored POS tagging mechanism to extract opinion words, as well as a Dutch semantic database to obtain sentiment scores for each opinion word (Li & Li, 2013). Sentiment polarity analysis includes attaching a positive or negative score to each product attribute mentioned. Using the product attributes that consumers mention in their pro and con sections as source for sentiment polarity, allows for fast and accurate scoring into positive (pro) and negative (con) scores as no opinion words need to be assessed. Furthermore, as indicated earlier, the pro and con sections of reviews generally provide a summary of the full text area (Branavan et al., 2009; Kim & Hovy, 2006). Hence, in this study, manual pro and con section analysis was considered the most effective and efficient approach for sentiment analysis as well.

Sentiment polarity was performed in the Excel datafile. For each unique attribute within the attribute category, a column was named and used for scoring the phrases referring to that attribute. Pro and con phrases of each review were read one by one, and sentiment polarity scores were directly assigned to the attribute referred to in the pro/con phrase. Based on Qi et al. (2016), attributes appearing in pro sections were scored with "1", those in con sections with "-1". Attributes that were not mentioned in a review, were labelled with a "0". This neutral score implies that the consumer has neither expressed a positive nor negative sentiment towards that attribute (Qi et al., 2016). A simplified example of the resulting data sheet is shown in Table 3.3; a section of the original table is included in Appendix H.

Table 3.3 Examples of attribute scores based on pro and con phrases

		Product attributes						
Pros	**Cons**	Coffee taste/quality	Speed of making coffee	Frequency of maintenance	Design (aesthetics)	Milk frother cleaning	Ease of operation	Coffee temperature
"delicious coffee within one minute"	*"requires maintenance frequently"*	1	1	-1				
"beautiful design"	*"cleaning after using the milk frother"*				1	-1		
"operating the machine is convenient"	*"the quality of the coffee is bad and temperature too low"*	-1					1	-1

Sentiment analysis was validated by a the same FAEM user who performed validation on attribute identification. Reliability of the validator as well as the number of reviews in the subset are described in section 3.3.1. The FAEM user was instructed to process the same subset of 85 reviews from the dataset, now featured with the final attribute list as columns. The validator was asked to transform the pro/con phrases of those reviews into attribute scores. Detailed instructions for the validation task, as were sent to the validator, are included in Appendix I. The results of the validation tasks were compared to the original document, which can be retrieved from Appendix J. Based on the 85 reviews; the attribute

scores assigned by the researcher and validator were compared based on the category scores that were computed by summing the scores of attributes within the category. 92.94% similarity was observed, which confirms inter-observer consistency as well as internal reliability of the sentiment analysis step. Furthermore, based on the correlations between the initial and validated attribute scores (amounting 1), the internal reliability (consistency) of the product attributes was ensured. (Bryman, 2015, p.170).

Based on the approach of Decker & Trusov (2010), attributes that were found to appear in less than 1% of all 565 reviews were reconsidered to avoid inaccurate outcomes. 1% is the empirically defined threshold of minimum support of a certain attribute in a review dataset to be relevant for further analysis (Ferreira et al., 2008; Hu & Liu, 2004[2]). In consultation with the same FAEM user (validator), those attributes and scores were either merged with another, related, product attribute or removed from the dataset if the attribute was found distinct and considered irrelevant to further analysis. Appendix A includes the final list of product attributes, grouped into capability, usability ant maintainability.

3.3.4 Computing variables in the model for statistical analysis

The final step in the data pre-processing was to prepare the variables for statistical analysis. The variables in the model had to be computed and imported to SPSS. The independent variable, number of features, was easily computed by adding a column to the dataset that contained the numbers 14, 16 and 21 for reviews of the low-, medium-, and high- product type, respectively. The dependent variable, consumer satisfaction, was also easily computed by retrieving the star ratings (1-5) that reviewers assigned as overall evaluation of the product (Decker & Trusov, 2010). The mediating variables, however, were to be computed from the set of scored product attributes within each category, resulting from step 3.3.3.

Preparing the mediating variables capability, usability, and maintainability for statistical analysis requires calculation of overall category scores for each product review. According to Hu and Liu (2004[1]), sentiment scores can be summed up to represent the aggregated sentiment of a higher-level category. Their approach was adopted for this final data pre-processing step. An Excel formula was used to sum the individual scores of all attributes belonging to the same category for each review (row). E.g., suppose a consumer mentioned coffee quality (capability), coffee temperature (capability) and ease of operation (usability) in the pro section, and frequency of cleaning (maintainability) in the con section. This consumer's overall category scores will be 2 for capability, 1 for usability and -1 for maintainability. The newly computed variables were imported in SPSS, after which the data analysis could start. Since this study also aims to report the specific reasons for the hypothesised mediating effects on consumer satisfaction, unique product attributes were also added to SPSS as distinct variables. Each attribute is later conceived as having a unique value and a unique effect on consumer satisfaction.

3.4 Data analysis

As can be seen from the proposed conceptual model, there is one independent variable, three parallel mediators, and one dependent variable. Based on the pre-processing steps described above, the variables in the model are computed based on different review data, as shown in Figure 3.3. The independent variable in the model is number of beneficial features, from now on referred to as "features". For each unique review, features is represented by the number of features of the FAEM model that is reviewed, i.e., either the low- (14), medium- (16), or high-feature (21) model. The dependent variable is consumer satisfaction, "satisfaction", represented by overall product score that is given by the reviewer, i.e., the number of stars (1 to 5). The three parallel mediators are perceived "capability", "usability", and "maintainability". Those are the sums of each category's attribute scores, as were retrieved based on the attributes' appearance in either the pro or con section of the review. Since each reviewer only used and

evaluated one out of three product types, no information is available on how consumers would review the other two products. Therefore, the data cannot be analysed at the consumer level. For each product type, i.e., each of three feature levels, consumer opinions are aggregated to represent the data at the feature level.

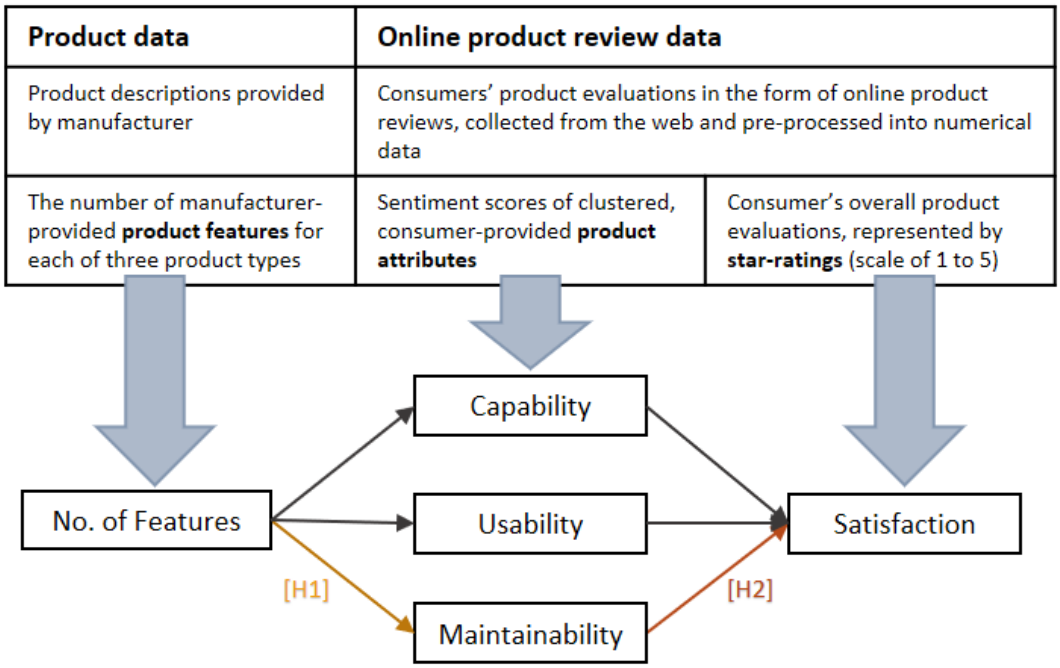

Figure 3.3: Conceptual model showing the data sources used to compute each variable

This research thus aims to transform individual consumer opinions into aggregate consumer needs for each level of features. Since this study extends the model developed by Thompson et al. (2005), a similar analysis should be appropriate to test and compare the relationships in the conceptual model. Thompson et al. (2005) used multi-sample path analysis to determine the relative effects of each relationship, because their before- and after-use measurements represent two distinct samples. Since this study only includes after-use measurement, a regular path analysis was used. Hayes (2012) developed an SPSS macro for path analysis of multiple mediator models. This macro enables comparison of the size of indirect effects through different mediators in one model (Hayes, 2012; MacKinnon et al., 2000). This is useful for competitive theory testing, as is the case in this study: it is aimed to compare the difference in the indirect effects through each mediator to see whether the extended model (i.e., including maintainability) complements the original model (Thompson et al., 2005) for maintenance-sensitive products. Specifically, the path analysis that is used to test the model is based on multiple linear regression with triple parallel mediation. Multiple linear regression measures the effect of more than one predictor on one dependent variable (Landau, 2004, p.99). Triple parallel mediation considers the role of capability, usability, and maintainability, that operate in parallel rather than in series, to transmit the effect between features and satisfaction. The mediators explain how or why this effect occurs (MacKinnon et al., 2002).

The multiple linear regression model with triple parallel mediation is equal to 'model 4' as specified by Hayes (2012; 2017) in the PROCESS macro for SPSS. The model 4-configuration is therefore selected to estimate the direct and indirect effects in the model. The direct effect of features (*F*) on satisfaction (*S*) is represented by equation 4.1 on the next page. The specific indirect effect of *F* on *S* through

mediator M_j is estimated as the product of a_j in equation 4.2 and b_j in equation 4.3 (Hayes & Rockwood, 2016).

These equations show that, the indirect effect of F on S through M_j, M_j should be regressed on F (equation 4.2), and S should be regressed on F and on all three mediators (equation 4.3). Mediator j has three forms: capability, usability, and maintainability. Hence, there are three versions of equation 4.3, and three a paths, one for each mediator. The a path of each mediator is multiplied by that mediator's corresponding effect on S in equation 4.3 (b_j) to get the specific indirect effect of F on S through M_j. The sum of the three specific indirect effects, when added to the direct effect of F (c' in equation 4.3) yields the total effect of F (c from equation 4.1) (Hayes & Rockwood, 2016). The PROCESS macro was used to calculate the direct and indirect effects in the model.

Eq. 4.1 $$\hat{S} = cF + e_Y$$

Eq. 4.2 $$\widehat{M}_j = a_j F + e_j$$

Eq. 4.3 $$\hat{S} = c'F + \sum_{j=1}^{3} b_j M_j + e_Y$$

4. Results

This chapter describes the results of the data analysis, that will provide an answer to the research question *"to what extent do perceived capability, usability, and maintainability of maintenance-sensitive products influence the effect of adding features on consumer satisfaction?"*. First, the hypotheses are tested and effects between the variables are reported and interpreted. Both partial effects as indirect effects of the number of features on consumer satisfaction are discussed. The path coefficients are mapped into the conceptual model to provide complete view of how the variables affect each other. Second, the relationships between individual product attributes and consumer satisfaction are reported and ranked based on size to provide guidelines for product improvements.

4.1 The effect of features on satisfaction, through capability, usability and maintainability

As indicated in the methodology, the statistical data analysis was performed using SPSS. The effect of number of beneficial features on consumer satisfaction was measured as direct effect, as well as indirect effect through perceived capability, usability and maintainability. The path analysis using the PROCESS macro in SPSS (Hayes, 2012) resulted in the standardised path coefficients shown in Figure 4.1. The direct effect of features on satisfaction is not significant ($p > 0.05$), but the indirect, partial effects through capability, usability and maintainability are. This indicates that the effect of features on satisfaction occurs through full mediation by at least one of the mediators. The effect of features on maintainability is negative and significant, indicating that when the number of beneficial features increases, consumers' perceptions of the product's maintainability decrease. This finding confirms Hypothesis 1. The effect of features on both capability and usability is positive and significant, indicating that when the number of beneficial features increases, consumers' perceptions of the product's capability and usability increase. Features seems to have the biggest effect on capability (0.113) followed by usability (0.091) and maintainability (-0.085). Notably, the effects of features on capability and usability are positive and significant, which was unexpected for usability based on Thompson et al. (2005). Furthermore, the effect of maintainability on satisfaction is positive and significant, confirming Hypothesis 2. When consumers perceive higher maintainability, their overall satisfaction with the product increases. Similar positive and significant effects are measured for capability and usability on satisfaction.

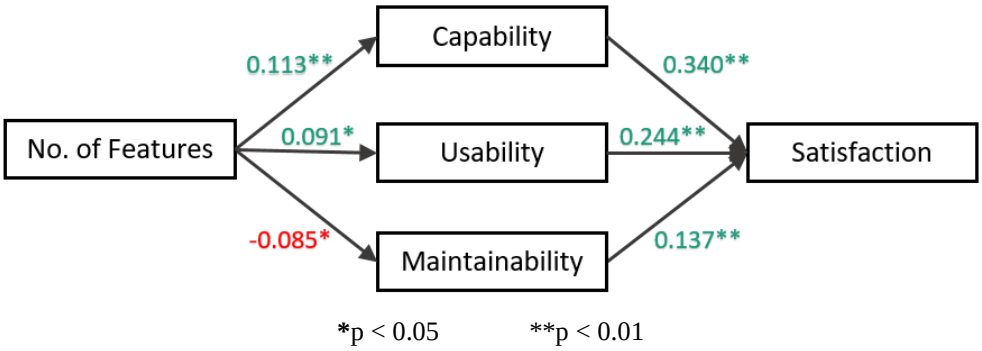

Figure 4.1: Standardised path coefficients resulting from path analysis

Table 4.2 includes the standardised indirect effects of features on satisfaction resulting from the path analysis. The beta for the mediating effect of maintainability is -0.0116, with a bootstrapped standard error of 0.0077, and a 95% confidence interval ranging from -0.0300 to 0.0011. Because this interval range includes zero, maintainability does not significantly mediate the relationship between features and satisfaction. Significant mediation does occur for capability and usability with betas of 0.0385 and 0.0222, respectively.

Table 4.1: Standardised indirect effects of Features on Satisfaction

	Effect	Boot SE	BootLLCI	BootULCI
TOTAL	0.0490	0.0190	0.0126	0.0875
Capability	0.0385	0.0150	0.0106	0.0708
Usability	0.0222	0.0097	0.0042	0.0430
Maintainability	-0.0116	0.0077	-0.0300	0.0011

4.2 The effect of individual product attributes on satisfaction

To find the attributes that have a significant effect on satisfaction, standardised effects of each unique product attribute were obtained for each FAEM type using multiple linear regression analysis of the attributes on satisfaction. Furthermore, the mean scores of the attributes were computed for each product type to represent consumers' evaluations of these attributes. The effect sizes and mean scores described in this paragraph can be used as guidelines for product improvements. It was not within the scope of this study to analyse product attributes at a fine-grained level. Therefore, the effect sizes and mean scores were not analysed further. Various approaches for further processing these attribute effects and mean scores into product-specific, tailored improvement strategies are available (Goswami & Tiwari, 2015; Jin et al., 2014; Qi et al., 2016; Li et al., 2014) that could be used for future extensions of this study.

The attributes that were found to have a significant effect on satisfaction are included in Appendix L. The effect sizes (betas) are mostly positive, indicating that an increase of the average consumer's perception of each product attribute (i.e., opinions become more positive) increases consumers' overall satisfaction with the product. This does also imply that when these attributes do not meet consumers' expectations and, hence, perceptions of the attribute are low (i.e., opinions are negative), satisfaction will decrease. Attributes that have a positive effect on consumer satisfaction, but were evaluated negatively thus provide the weaknesses of current FAEMs. Table 4.2 includes an overview of current weaknesses of the 14- and 16-feature product, ranked based on the effect on consumer satisfaction. For the 21-feature product, no mean attribute evaluations were negative.

Table 4.2: Product attributes that have negative mean scores, ranked based on effect on satisfaction

HD8821 (14 features)	HD8824 (16 feature)
1. Milk frother performance/quality	1. Milk frother sound
2. Ease-of-use	2. Ease-of-cleaning
3. Machine size	3. Milk frother performance/quality
4. Grinder sound	
5. Bean container ease-of-use	
6. Machine sound (in general)	

5. Discussion

Using online product reviews as source for identifying consumer needs, this study aimed to measure the effect of adding product features on consumer satisfaction, through consumers' perceptions of capability, usability and maintainability of the product. The focus of this study is maintenance-sensitive products, for which maintainability can be problematic as the complexity of the product, i.e., number of features, increases. For this type of products, maintainability operates as additional factor of consumer satisfaction. This study therefore extended the model of Thompson et al. (2005) by adding perceived maintainability as third mediator, in addition to perceived capability and perceived usability, of the relationship between product features and consumer satisfaction. A case study approach was used, based on Philips' FAEMs. Philips recognises increasing consumer needs for maintainability, as their FAEMs become more and more technologically advanced and complex. However, the key insights that were missing are how consumers balance their needs for capability, usability and maintainability, and how these change for different feature levels. This study provides product developers with these crucial insights, helping them to identify product improvements that will increase future consumer satisfaction and new product success. The trend of adding product features, possibly causing feature fatigue, was represented by three types of Philips' FAEMs, each having a different number of product features. To integrate the voice of the consumer in NPD, product attributes, opinions about these attributes and the sentiment polarity of these opinions were extracted from consumers' product evaluations in the form of online product reviews. The resulting attributes and scores were clustered into capability, usability, and maintainability, representing the underlying consumer needs that determine consumer satisfaction. Consumers' star ratings of the overall product, on a scale from 1 to 5, were used as measure of consumer satisfaction. The effects in the model were estimated through multiple linear regression-based path analysis using the PROCESS macro in SPSS. This led to important findings, which are discussed below.

The most important findings of this study include the significant, partial effects of features on maintainability, and in turn, maintainability on satisfaction. Hypothesis 1 was confirmed, since features was found to have a negative (-0.085), significant ($p<0.05$) effect on maintainability. When the number of features in a product increases, consumers perceive the maintainability of the product more negatively. While -0.085 would initially not seem like a big impact; its actual meaning should be noted: for each feature that is added to a FAEM, consumers' perceived maintainability decreases with 8.5%. Hypothetically, when comparing the € 400-worth entry-model FAEM with 14 features to the € 600-worth higher-end model with 21 features, consumers would be charged an additional € 200 for a model that they will eventually perceive as ($1-0.915^7=$) 46.30% less maintainable. Furthermore, hypothesis 2 was also confirmed, since maintainability was found to have a positive (0.137), significant ($p<0.01$) effect on satisfaction. For each unit increase in perceived maintainability, consumer satisfaction increases with 13.7%. In the previous hypothetical example, isolating the partial effects of maintainability from the effects of capability and usability, satisfaction would be decreased by ($53.70*0.863=$) 53.7% for the 21-feature model compared to the 14-feature model due to decreased maintainability. These findings indicate that perceived maintainability in fact plays an important role in the relationship between adding features and consumer satisfaction for maintenance-sensitive products.

Perceived maintainability should however be interpreted relative to consumers' perceptions of capability and usability within the features-satisfaction relationship. Specifically, as emphasised in the theoretical background of this study, consumers balance their needs for capability, usability and maintainability when evaluating a maintenance-sensitive product. For this reason, in addition to the partial effects described above, the combined indirect effects of features on satisfaction, through capability, usability and maintainability were measured. The path analysis revealed an effect size of -0.012 for the mediating effect of maintainability, which, against expectations, proved insignificant ($p>0.05$). Although partial

effects of features on maintainability and maintainability on satisfaction were significant, it should thus be noted that maintainability does not significantly mediate the relationship between features and satisfaction in the complete model. This implies that, for this case study, increasing product features does not significantly affect consumer satisfaction through consumers' decreased perceptions of maintainability. Significant mediation does occur for capability and usability with betas of 0.039 and 0.022, respectively. Hence, increasing product features does significantly increase consumer satisfaction through both capability and usability, which provides the answer to the research question "*To what extent do perceived capability, usability, and maintainability of maintenance-sensitive products influence the effect of adding features on consumer satisfaction?".* A probable explanation for the insignificance of the mediating effect of maintainability might become more necessary and frequent after longer periods of use. Specifically, in the case of Philips FAEMs, several maintenance actions might not be required within the first two months, depending on water hardness and frequency of using the machine (Philips, 2017). Since the online reviews in the database of this study were generally posted within 1.38 months of usage, maintenance issues might not have been discovered yet and are therefore less likely to be evaluated in the review. It is therefore expected that, over time, maintainability will start to operate as significant mediator in the features-satisfaction relationship.

Comparing the results of this study to the results of Thompson et al. (2005), also reveals notable findings. In both studies, the indirect, partial effects measured for capability and usability on the features-satisfaction relationship are significant. In this study, the effect of features on both capability (0.113) and usability (0.091) is positive and significant ($p<0.01$ for capability, $p<0.05$ for usability). This implies that when the number of features increases, consumers' perceptions of the product's capability and usability increase. The positive effect of features on usability is however in contradiction with the predictions and findings of Thompson et al. (2005) on the feature fatigue effect. The feature fatigue effect implies that adding product features increases satisfaction through increased capability, but at the same time, decreases satisfaction through decreased usability of the product. Since in the current study perceived usability increases with the number of features, it can be stated that the feature fatigue effect does not occur for Philips' FAEMs product line. Reasonable explanations for this contrasting effect on usability include the difference in product types studied, and Philips' customer-focused and user-friendly focus, as mentioned in the theoretical background. Firstly, FAEMs differ from digital audio and video players (Thompson et al., 2005) in many aspects, such as tangibility, consumers' status quo, output delivered, usage context, life cycle duration and price. These fundamental differences infer divergent consumer needs as well as expectations of the product's capability and usability. Secondly, Philips has been focusing increasingly on creating products that are not only advanced in terms of technology, but customer-focused and user-friendly as well, implied by their former slogan 'Sense and Simplicity' (NOS, 2013). This implies that the importance of product usability could already be accounted for sufficiently in Philips' NPD, which was also one of Thompson et al.'s (2005) recommendations to product developers to defeat feature fatigue. When feature fatigue has already been defeated, it does not mean that a product's perceived capability, usability and maintainability are optimised. There are still many challenges ahead, for which specific recommendations are provided in the conclusion.

6. Conclusion

The findings of this study, as described in the discussion, can be translated into theoretical contributions and managerial implications, which are discussed in this concluding chapter. The conclusion ends with limitations of this study and directions for future research.

The theoretical contributions of this study are widely applicable, as this study connects to multiple research disciplines, including NPD, product design, innovation, consumer behaviour, marketing, management and online product review literature. The most relevant contributions are discussed, starting from three perspectives: the role of maintainability in consumer research, the use of online product reviews as source for consumer research, and the operationalisation of capability, usability and maintainability as bundle of product attributes.

Firstly, this study closes the literature gap of the role of maintainability in relation to consumer satisfaction for maintenance-sensitive, physical products. Perceived maintainability was therefore assessed as third underlying consumer need, in addition to capability and usability (Thompson et al., 2005), mediating the relationship between product features and consumer satisfaction. The resulting findings confirm the hypothesised influence of adding features on perceived maintainability and importance of maintainability to consumer satisfaction. These findings extend literature on the feature fatigue effect (e.g., Thompson et al., 2005; Rust et al., 2006); to also be applicable to maintenance-sensitive products. Furthermore, in NPD and product design literature, the relation between product complexity and maintainability for software products (Slavila et al., 2005) can now be generalised to physical products. In a similar way, usability research, originating from software studies, has become a crucial part of physical product research (Han et al., 2001). The connection established in this study fosters the integration of insights from software-related research into physical product research in NPD, design and development literature. For instance, consumer research for the development of maintenance-sensitive products should account for consumer needs for maintainability when establishing engineering requirements. Furthermore, as this study confirms the influence of maintainability on satisfaction for maintenance-sensitive products, it also contributes to literature on consumer behaviour and marketing by introducing maintainability as new perspective to create value to consumers.

Secondly, by using online product reviews as data source, this study expands the range of consumer research literature for NPD and marketing applications. Rather than using experimental settings, trial usage, surveys and simulated consumers (e.g., Thompson et al., 2005), this study is based on real usage situations of actual consumers as reflected by online product reviews (Srinivasan et al., 1997). As discussed earlier, online reviews offer a promising new information source, allowing for more reliable, effective and efficient consumer research for involving the VOC in NPD. This study can be seen as another successful application of online review data for NPD and marketing purposes, and therefore serve as an example to scholars that are willing to exploit the wealth of information embedded in freely accessible online consumer evaluations.

Thirdly, this study contributes to several research disciplines by the operationalisation of capability, usability and maintainability as bundle of product attributes. Previous online product review research has already shown to effectively predict consumer satisfaction from consumers' evaluations of product attributes extracted from online reviews (e.g., Decker & Trusov, 2010; Qi et al., 2016, Wang & Wang, 2014). The results of this study demonstrate that also underlying consumer needs, such as capability, usability and maintainability can be successfully extracted from online product reviews to predict consumer satisfaction. This helps scholars to frame existing concepts in literature in new ways and develop product improvements on strategy-level. Furthermore, insights based on underlying consumer

need categories could enhance understanding consumers' product evaluation processes. Consumers tend to structure their knowledge of specific product alternatives in categories as cognitive process (Gutman, 1982; Yang, 2013). Assessing consumers' evaluations of product attributes in categorisations, might therefore enhance understanding of why consumers prefer certain attributes over others and how this impacts consumer satisfaction as well was buying behaviour.

On top of these theoretical contributions, this study provides relevant managerial implications. An important part of NPD is to translate consumer needs into engineering requirements. Product developers often have extensive engineering knowledge, but lack insights from the consumers point of view. As discussed earlier, integration of the VOC in NPD is crucial to develop a product that actually satisfies consumers and therefore maximises chances of product success in the market. This study helps product developers of maintenance-sensitive products to understand manifest as well as latent consumer needs for capability, usability and maintainability from a usage-perspective. More specifically, it provides insights into how these needs are balanced, influencing the relationship between adding features and consumer satisfaction. The next two sections provide generalising recommendations for firms operating in industries of maintenance-sensitive products, as well as specific recommendations for Philips regarding their FAEMs.

At large, the findings of this study demonstrate that maintainability is an important consumer need, which becomes problematic when the number of product features increases. Furthermore, maintainability was found to positively influence consumer satisfaction. Product developers of maintenance-sensitive products are therefore recommended to account for maintainability as design requirement, especially when products become increasingly technologically advanced. However, needs for maintainability should be balanced against needs for capability and usability. In absence of priority information, product developers may instead address development objectives that risk new product's consumer satisfaction and market success.

In the specific context of the case study that was applied, Philips FAEMs, maintainability should not be the focal point of product development, but should not be neglected neither. The path analysis of the complete conceptual model revealed that the mediating effect of maintainability is not significant when accounted for the significant mediating effects of capability and usability. Based on the indirect effects that were measured, Philips is recommended to focus primarily on capability and secondarily on usability when extra features are added to new versions of FAEMs. The insignificant indirect effect of maintainability does not imply that no attention should be paid to maintenance-aspects at all. In fact, when looking at the partial effects in the model, it shows that when the number of features increases, consumers' perceptions of maintainability decrease significantly. As mentioned in the discussion, decreased maintainability for higher-level, more expensive FAEMs is in contrast with consumers' product expectations. To better manage consumer expectations, maintainability should be accounted for in NPD, in addition to capability and usability.

This study has several limitations, which provide directions for future research. The first limitation is the time frame in which product reviews are posted after the moment of purchase. As reported in the methodology section, the average time consumers own a FAEM when posting a review is 1.38 months. This might imply that some maintenance operations have not yet been required at the moment of writing a review. For instance, on their web page, Philips indicates that lubricating the brewing mechanism and descaling might not be needed until up to 2-6 months after purchase, dependent on usage frequency and regional water hardness (Philips, 2017). The current study might therefore have been unable to capture consumer evaluations of the complete spectrum of maintainability. Future research could therefore focus on how consumers' perceptions of maintainability change over time. Additionally, it would be relevant to examine whether weight of capability, usability and maintainability change over time, as well as their

effect on the features-satisfaction relationship. One way to obtain data which is comparable to the type of data used in this study, is asking registered product owners to write a review after several months or even years of usage.

The second limitation of the study is the scalability of the approach. Since many steps in the data pre-processing stage had to be performed manually, the scalability of this method is highly dependent on the time and human resources available. When a firm like Philips would decide to conduct online review analysis on big scale or on a regular basis, it is recommended to invest in acquiring or developing appropriate (semi-)automatic tools for data crawling, attribute identification, opinion extraction and sentiment analysis. A relevant direction for future research would therefore be to compare the available NLP-based tools for the purpose of improving products based on online product reviews. Those tools could be compared based on accuracy, cost, applicability to different review formats, applicability to less common languages (e.g., Dutch), degree of human intervention required, degree of model training required.

The third limitation of the study is the measure used for consumer satisfaction. Although the overall product rating of reviews as measure for satisfaction is motivated and often applied in previous research, consumers' interpretations of this scale differ (Qi et al., 2016). For instance, consumers might give three, four or five stars when they feel satisfied with a product. Furthermore, in this particular study, satisfaction scores are skewed to the right-tale, making it hard to distinguish different degrees of consumer satisfaction. Currently, no overview is available that compares different measures of consumer satisfaction that could be used for online review analysis. Therefore, future research could focus on comparing available measures and developing an optimal method to reflect differences in consumer satisfaction.

Other relevant directions of future research in a similar research setting include the identification of consumer segments to account for heterogeneity of consumer preferences, and the comparison of different types of maintenance-sensitive products. In essence, consumer research should keep up with the accelerating growth of online data. In addition to many existing relevant research on the analysis of online data, this study provides another starting point for the development of tools and methods to automatically obtain market and consumer insights in real-time.

7. References

Abraham-Murali, L., & Littrell, M.A. (1995). Consumers' conceptualization of apparel attributes. *Clothing and Textiles Research Journal*, *13*(2), 65-74.

Abrahams, A.S., Jiao, J., Wang, G.A., & Fan, W. (2012). Vehicle defect discovery from social media. *Decision Support Systems*, *54*(1), 87-97.

Ante, S.E. (2009). Amazon: turning consumer opinions into gold. Retrieved at March 15, 2017 from: www.businessweek.com/print/magazine/content/09_43/

Archak, N., Ghose, A., & Ipeirotis, P.G. (2011). *Deriving the pricing power of product features by mining consumer reviews. Management science*, 57(8), 1485-1509.

Bickart, B., & Schindler, R.M. (2001). Internet forums as influential sources of consumer information. *Journal of interactive marketing*, *15*(3), 31-40.

Blythe, J. (2008). *Essentials of marketing, 5th edition*. FT Prentice-Hall (Pearson Education), Harlow.

Boehm, B.W., Brown, J.R., & Lipow, M. (1976, October). Quantitative evaluation of software quality. In *Proceedings of the 2nd international conference on Software engineering* (pp. 592-605). IEEE Computer Society Press.

Bol.com (2018). Jouw mening over producten. Retrieved at January 24, 2018 from: https://www.bol.com/nl/m/klantenservice/productreviews/index.html

Branavan, S.R.K., Chen, H., Eisenstein, J., & Barzilay, R. (2009). Learning document-level semantic properties from free-text annotations. Journal of Artificial Intelligence Research, 34, 569-603.

Brown, C.L., & Carpenter, G.S. (2000). Why is the trivial important? A reasons-based account for the effects of trivial attributes on choice. *Journal of Consumer Research*, *26*(4), 372-385.

Bruce, R.F., & Wiebe, J.M. (1999). Recognizing subjectivity: a case study in manual tagging. *Natural Language Engineering*, *5*(2), 187-205.

Bryman, A. (2015). *Social research methods.* Oxford university press.

Burke, P.F. (2013). Seeking Simplicity in Complexity: The Relative Value of Ease of Use (EOU)-Based Product Differentiation. *Journal of Product Innovation Management*, *30*(6), 1227-1241.

BusinessWire (2016). Global Automatic Espresso Machines Market to Record Persistent Growth Due to Growing Demand from Western Europe and North America until 2019, Says Technavio. Retrieved at December 15, 2016 from: http://www.businesswire.com/news/home/20160218005045/en/Global-Automatic-Espresso-Machines-Market-Record-Persistent

Chen, Y., & Xie, J. (2008). Online consumer review: Word-of-mouth as a new element of marketing communication mix. *Management science*, *54*(3), 477-491.

Chen, Y., Fay, S., & Wang, Q. (2011). The role of marketing in social media: How online consumer reviews evolve. *Journal of Interactive Marketing, 25*(2), 85-94.

Cheung, C.M., & Thadani, D.R. (2012). The impact of electronic word-of-mouth communication: A literature analysis and integrative model. *Decision support systems, 54*(1), 461-470.

Chevalier, J.A., & Mayzlin, D. (2006). The effect of word of mouth on sales: Online book reviews. *Journal of marketing research, 43*(3), 345-354.

Christensen, C.M., Suárez, F.F., & Utterback, J. M. (1998). Strategies for survival in fast-changing industries. *Management science, 44*(12-part-2), S207-S220.

Clemons, E.K., Gao, G.G., & Hitt, L.M. (2006). When online reviews meet hyperdifferentiation: A study of the craft beer industry. *Journal of management information systems, 23*(2), 149-171.

Collobert, R., & Weston, J. (2008). A unified architecture for natural language processing: Deep neural networks with multitask learning. *Proceedings of the 25th international conference on Machine learning*. 160-167. ACM.

Coolblue (2017). Philips 3000 HD8821/01 Reviews. Retrieved on October 14, 2017 from: https://www.coolblue.nl/product-reviews/562411/philips-3000-hd8821-01.html

Coolblue (2018). Richtlijnen voor reviews. Retrieved at January 28, 2018 from: https://www.coolblue.nl/klantenservice/1441/informatie-over/richtlijnen-voor-reviews.html

Cooper, R.G., Edgett, S.J., & Kleinschmidt, E.J. (2004). Benchmarking best NPD practices—II. *Research-Technology Management, 47*(3), 50-59.

Das, S.R., & Chen, M.Y. (2007). Yahoo! for Amazon: Sentiment extraction from small talk on the web. *Management science, 53*(9), 1375-1388.

Dash, N.S. (2013). Part-of-Speech (POS) Tagging in Bengali Written Text Corpus. *International Journal on Linguistics and Language Technology, 1*(1), 53-96.

De Maeyer, P. (2012). Impact of online consumer reviews on sales and price strategies: A review and directions for future research. *Journal of Product & Brand Management, 21*(2), 132-139.

Decker, R., & Trusov, M. (2010). Estimating aggregate consumer preferences from online product reviews. *International Journal of Research in Marketing, 27*(4), 293-307.

Ferreira, L., Jakob, N., & Gurevych, I. (2008, August). A comparative study of feature extraction algorithms in customer reviews. In *Semantic Computing, 2008 IEEE International Conference on* (pp. 144-151). IEEE.

Filieri, R. (2015). What makes online reviews helpful? A diagnosticity-adoption framework to explain informational and normative influences in e-WOM. *Journal of Business Research, 68*(6), 1261-1270.

Fornell, C. (1992). A national customer satisfaction barometer: The Swedish experience. *The Journal of Marketing*, 6-21.

Frizziero, L. (2014). A coffee machine design project through innovative methods: QFD, value analysis and design for assembly. *Asian Research Publishing Network (ARPN). Journal of Engineering and Applied Sciences. ISSN*, *6608*(9), 7.

Garrette, D., Mielens, J., & Baldridge, J. (2013). Real-World Semi-Supervised Learning of POS-Taggers for Low-Resource Languages. In *ACL (1)* (pp. 583-592).

Ghose, A., & Ipeirotis, P.G. (2007). Designing novel review ranking systems: predicting the usefulness and impact of reviews. *Proceedings of the ninth international conference on Electronic commerce*, 303-310.

Goldenberg, J., Horowitz, R., Levav, A., & Mazursky, D. (2003). Finding your innovation sweet spot. *Harvard Business Review*, *81*(3), 120-129.

Goodman, J.K., & Irmak, C. (2013). Having versus consuming: failure to estimate usage frequency makes consumers prefer multifeature products. *Journal of Marketing Research*, *50*(1), 44-54.

Goswami, M., & Tiwari, M.K. (2015). Product feature and functionality driven integrated framework for product commercialization in presence of qualitative consumer reviews. *International Journal of Production Research*, *53*(16), 4769-4788.

Griffioen A. (2008). *Het Senseo-effect*, Thema.

Gutman, J. (1982). A means-end chain model based on consumer categorization processes. *The Journal of Marketing*, 60-72.

Gwet, K.L. (2016). Testing the difference of correlated agreement coefficients for statistical significance. *Educational and Psychological Measurement*, *76*(4), 609-637.

Han, S.H., Yun, M.H., Kwahk, J., & Hong, S.W. (2001). Usability of consumer electronic products. *International Journal of Industrial Ergonomics*, *28*(3-4), 143-151.

Hardey, M. (2011). Generation C: content, creation, connections and choice. *International Journal of Market Research*, *53*(6), 749-770.

Hayes, A.F. (2012). PROCESS: A versatile computational tool for observed variable mediation, moderation, and conditional process modeling.

Hayes, A.F. (2013). *Methodology in the social sciences. Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. New York: Guilford Press.

Hayes, A.F., & Rockwood, N.J. (2016). Regression-based statistical mediation and moderation analysis in clinical research: Observations, recommendations, and implementation. *Behaviour research and therapy*.

Hayes, A.F. (2017). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach.* Guilford Publications.

Hooks, I.F., & Farry, K.A. (2001). *Customer-centered products: creating successful products through smart requirements management.* AMACOM Div American Mgmt Assn. ISO 690.

Hu, M., & Liu, B. (2004[1], August). Mining and summarizing customer reviews. In *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 168-177). ACM.

Hu, M., & Liu, B. (2004[2], July). Mining opinion features in customer reviews. In *AAAI* (Vol. 4, No. 4, pp. 755-760).

Hu, N., Pavlou, P.A., & Zhang, J. (2006, June). Can online reviews reveal a product's true quality?: empirical findings and analytical modeling of Online word-of-mouth communication. In *Proceedings of the 7th ACM conference on Electronic commerce* (pp. 324-330). ACM.

Huang, J.H., & Chen, Y.F. (2006). Herding in online product choice. *Psychology & Marketing*, *23*(5), 413-428.

ISO/IEC (1996). *Information Technology - Software Quality characteristics and metrics - Part 1: Quality characteristics and sub-characteristics*, pp. 21.

Jin, J., Ji, P., & Gu, R. (2016). Identifying comparative customer requirements from product online reviews for competitor analysis. *Engineering Applications of Artificial Intelligence*, *49*, 61-73.

Jin, J., Ji, P., & Liu, Y. (2014). Prioritising engineering characteristics based on customer online reviews for quality function deployment. *Journal of Engineering Design*, *25*(7-9), 303-324.
Jindal, N., & Liu, B. (2008, February). Opinion spam and analysis. In *Proceedings of the 2008 International Conference on Web Search and Data Mining* (pp. 219-230). ACM.

Kangale, A., Kumar, S.K., Naeem, M.A., Williams, M., & Tiwari, M.K. (2016). Mining consumer reviews to generate ratings of different product attributes while producing feature-based review-summary. *International Journal of Systems Science*, *47*(13), 3272-3286.

Karwowski, W., Soares, M.M., & Stanton, N.A. (Eds.). (2011). *Human factors and ergonomics in consumer product design: Uses and Applications*. CRC Press.

Kekre, S., Krishnan, M.S., & Srinivasan, K. (1995). Drivers of customer satisfaction for software products: implications for design and service support. *Management science*, *41*(9), 1456-1470.

Khurana, A., & Rosenthal, S.R. (1998). Towards holistic "front ends" in new product development. *Journal of product innovation management*, *15*(1), 57-74.

Kieskeurig (2018). Reviews. Retrieved at January 28, 2018 from: https://www.kieskeurig.nl/reviews.

Kim, S.M., & Hovy, E. (2006). Automatic identification of pro and con reasons in online reviews. *Proceedings of the COLING/ACL on Main conference poster sessions*. 483-490. Association for Computational Linguistics.

Kotler, P., & Armstrong, G. (2010). *Principles of marketing*. Pearson education.

Kotler, P., Keller, K.L. (2006). *Marketing management*, 12th Edition, Pearson Education Inc, New Jersey

Landau, S. (2004). *A handbook of statistical analyses using SPSS*. CRC.

Lee, T., & Bradlow, E.T. (2007). Automatic construction of conjoint attributes and levels from online customer reviews. *University of Pennsylvania, The Wharton School Working Paper*.

Li, X., Hitt, L.M., & Zhang, Z.J. (2011). Product reviews and competition in markets for repeat purchase products. *Journal of Management Information Systems*, *27*(4), 9-42.

Li, Y.M., & Li, T.Y. (2013). Deriving market intelligence from microblogs. *Decision Support Systems*, *55*(1), 206-217.

Liu, B., Hu, M., & Cheng, J. (2005). Opinion observer: analyzing and comparing opinions on the web. *Proceedings of the 14th international conference on World Wide Web*, 342-351.

Lok, C.Y., Mazlina, M.S., & Tuah, B.H. (2013). Development of an integrated grating and slicing machine for starchy vegetables. *Journal of Food, Agriculture & Environment*, *11*(1), 141-145.

MacKinnon, D. P., Krull, J. L., & Lockwood, C. M. (2000). Equivalence of the mediation, confounding and suppression effect. *Prevention science*, *1*(4), 173-181.

MacKinnon, D.P., Lockwood, C.M., Hoffman, J.M., West, S.G., & Sheets, V. (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychological methods*, *7*(1), 83.

Mari, M. (2003, September). The impact of maintainability on component-based software systems. In *null* (p. 25). IEEE.

Mayzlin, D. (2006). Promotional chat on the Internet. *Marketing Science*, *25*(2), 155-163.

Mayzlin, D., Dover, Y., & Chevalier, J. (2014). Promotional reviews: An empirical investigation of online review manipulation. *American Economic Review*, *104*(8), 2421-55.

Mick, D.G. & Fournier, S. (1998). Paradoxes of Technology: Consumer Cognizance, Emotions, and Coping Strategies. *Journal of Consumer Research*, *25*(9), 123-43.

Morris, J. (2013). Why espresso? Explaining changes in European coffee preferences from a production of culture perspective. *European Review of History: Revue européenne d'histoire*, *20*(5), 881-901.

Mukherjee, A., & Hoyer, W.D. (2001). The effect of novel attributes on product evaluation. *Journal of Consumer Research, 28*(3), 462-472.

Netzer, O., Feldman, R., Goldenberg, J., & Fresko, M. (2012). Mine your own business: Market-structure surveillance through text mining. *Marketing Science, 31*(3), 521-543.

Netzer, O., Toubia, O., Bradlow, E.T., Dahan, E., Evgeniou, T., Feinberg, F.M., Orlin, J.B. (2008). Beyond conjoint analysis: Advances in preference measurement. *Marketing Letters, 19*(3-4), 337-354.

Nielsen, J. (1994). *Usability engineering.* Elsevier.

Nielsen, J., & Landauer, T.K. (1993, May). A mathematical model of the finding of usability problems. In *Proceedings of the INTERACT'93 and CHI'93 conference on Human factors in computing systems* (pp. 206-213). ACM.

NOS (2013). Philips wil meer mensen bereiken. Retrieved at January 27, 2018 from: https://nos.nl/artikel/574670-philips-wil-meer-mensen-bereiken.html.

Nowlis, S.M., & Simonson, I. (1996). The effect of new product features on brand choice. *Journal of marketing research*, 36-46.

Pang, B., Lee, L., & Vaithyanathan, S. (2002). Thumbs up?: sentiment classification using machine learning techniques. *Proceedings of the ACL-02 conference on Empirical methods in natural language processing-Volume 10*, 79-86.

Pecht, M. (Ed.). (2009). *Product reliability, maintainability, and supportability handbook*. CRC Press.

Perez, S. (2016). Amazon makes good on its promise to delete "incentivized" reviews. Retrieved at Jauary 25, 2018 from: https://techcrunch.com/2016/11/23/amazon-made-good-on-its-promise-to-delete-incentivized-reviews/

Petrescu, M., O'Leary, K., Goldring, D., & Mrad, S.B. (2017). Incentivized reviews: Promising the moon for a few stars. *Journal of Retailing and Consumer Services*.

Philips (2017). Cleaning and maintaining your Philips espresso machine. Retrieved at November 11, 2017 from: https://www.usa.philips.com/c-m-ho/coffee/philips-espresso-machine-care.

Porter, M.E. (1985). *Competitive Advantage: Creating and Sustaining Superior Performance*. Free Press, New York.

Porter, M.E. (1998). *Clusters and the new economics of competition* (Vol. 76, No. 6, pp. 77-90). Boston: Harvard Business Review.

Qi, J., Zhang, Z., Jeon, S., & Zhou, Y. (2016). Mining customer requirements from online reviews: A product improvement perspective. *Information & Management, 53*(8), 951-963.

Reimer, T., & Benkenstein, M. (2016). Altruistic eWOM marketing: More than an alternative to monetary incentives. *Journal of Retailing and Consumer Services, 31*, 323-333.

Rivlin, G. (2005). Founders of Web Site Accuse Backers of Cheating Them. Retrieved on December 20, 2017 from: http://www.nytimes.com/2005/01/26/technology/founders-of-web-site-accuse-backers-of-cheating-them.html

Rust, R.T., Thompson, D.V., & Hamilton, R.W. (2006). Defeating feature fatigue. *Harvard business review, 84*(2), 37-47.

Sauro, J., & Lewis, J. R. (2016). *Quantifying the user experience: Practical statistics for user research.* Morgan Kaufmann. ISO 690

Seliger, G., Zussman, E., & Kriwet, A. (1994). Integration of recycling considerations into product design—A system approach. In *Information and Collaboration Models of Integration* (pp. 27-41). Springer, Dordrecht.

Seva, R.R., Helander, M.G., & Duh, H.B.L. (2007). Integrating pre-purchase affect in product concept development. *International Journal of Simulation and Process Modelling, 3*(4), 195-203.

Scott, W.A. (1955). Reliability of content analysis: The case of nominal scale coding. *Public opinion quarterly,* 321-325.

Shimomura, Y., Tanigawa, S., Umeda, Y., & Tomiyama, T. (1995). Development of self-maintenance photocopiers. *AI Magazine*, *16*(4), 41.

Slavila, C.A., Decreuse, C., & Ferney, M. (2005). Fuzzy approach for maintainability evaluation in the design process. *Concurrent Engineering, 13*(4), 291-300.

Srinivasan, V., Lovejoy, W.S., & Beach, D. (1997). Integrated product design for marketability and manufacturing. *Journal of Marketing Research*, 154-163.

Stake, R.E. (2000). Case Studies In: Denzin N. and Lincoln, Y.(Ed.). *Handbook of qualitative research* (pp. 435-455).

Tal-Or, N., Cohen, J., Tsfati, Y., & Gunther, A.C. (2010). Testing causal direction in the influence of presumed media influence. *Communication Research, 37*(6), 801-824.

Thelwall, M., Buckley, K., Paltoglou, G., Cai, D., & Kappas, A. (2010). Sentiment strength detection in short informal text. Journal of the Association for Information Science and Technology, 61(12), 2544-2558.

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.

Thompson, D.V., & Norton, M.I. (2011). The social utility of feature creep. *Journal of Marketing Research, 48*(3), 555-565.

Tse, D.K., & Wilton, P.C. (1988). Models of consumer satisfaction formation: An extension. *Journal of marketing research*, 204-212.

Van Kralingen, R. (2007). *De groeimotor*. Pearson Education.

Wang, H., & Wang, W. (2014). Product weakness finder: an opinion-aware system through sentiment analysis. *Industrial Management & Data Systems*, *114*(8), 1301-1320.

Wiklund, M.E. (1994). *Usability in practice: how companies develop user-friendly products.* Boston: Academic Press.

Wu, M., Wang, L., Li, M., & Long, H. (2014). An approach of product usability evaluation based on Web mining in feature fatigue analysis. *Computers & Industrial Engineering*, *75*, 230-238.

Xiao, S., Wei, C.P., & Dong, M. (2016). Crowd intelligence: Analyzing online product reviews for preference measurement. *Information & Management*, *53*(2), 169-182.

Xu, Z., Frankwick, G.L., & Ramirez, E. (2016). Effects of big data analytics and traditional marketing analytics on new product success: A knowledge fusion perspective. *Journal of Business Research*, *69*(5), 1562-1566.

Yang, S. (2013, March). Consumers' evaluation of brands from emerging countries. In *RobertMmittelstaedt doctoral symposium proceedings* (p. 103).

Yin, R.K. (1981). The case study crisis: Some answers. *Administrative science quarterly*, *26*(1), 58-65.

Yin, R.K. (2013). *Case study research: Design and methods.* Sage publications.

Zhou, W., & Duan, W. (2015). An empirical study of how third-party websites influence the feedback mechanism between online word-of-mouth and retail sales. *Decision Support Systems, 76*, 14-23.

Zhu, F., & Zhang, X. (2010). Impact of online consumer reviews on sales: The moderating role of product and consumer characteristics. *Journal of marketing*, *74*(2), 133-148.

Appendix A: Features of the FAEMs of Philips that were studied

	HD8821	HD8824	HD8834
			
Feature level	Low	Medium	High
Total no. of features	14	16	21
User profiles	✓	✓	✓
Coffee volume adjustability	✓	✓	✓
Milk volume adjustability	✗	✗	✓
Aroma strength adjustability	✗	✗	✓
Grinder settings	✓	✓	✓
Temperature adjustability	✗	✗	✓
LED buttons	✓	✓	✓
Basic display	✗	✗	✓
Milk frother function*	✓	✓	✓
Removable brewing group	✓	✓	✓
Automatic cleaning and guided descaling	✓	✓	✓
Fast heating boiler	✓	✓	✓
Water filter (Brita) compatibility	✓	✓	✓
Espresso function**	✓	✓	✓
Espresso lungo function**	✗	✓	✓
Hot water function**	✓	✓	✓
Cappuccino function**	✗	✓	✓
Coffee function**	✓	✓	✓
Double cup function	✓	✓	✓
Ground coffee function	✗	✗	✓
* Different milk frothers:	Panarello	Cappuccinatore	Automatic; w/ carafe

** Drinks by one push on the button (fully automatically)

Appendix B: Review characteristics and reviewer demographics

Source	Pr.Type	No. reviews	Average age	Sex (M : F)	Average time use (months)	Pro:Con	Recomm. (Y : N)	Average satisfaction
Philips.nl	HD8821	19	45.21	13 : 6	2.37	2.89 : 1.89	11 : 8	3.58
	HD8824	10	45.85	6 : 4	0.65	2.20 : 0.90	10 : 0	4.60
	HD8834	3	52.17	3 : 1	1.00	1.00 : 1.00	2 : 1	3.00
	Total	32	46.06	22 : 11	1.70	2.50 : 1.50	23 : 9	3.84
Bol.com	HD8821	35	47.67	-	-	2.80 : 0.69	32 : 3	4.23
	HD8824	30	39.95	-	-	3.07 : 0.53	29 : 1	4.60
	HD8834	8	46.00	-	-	2.50 : 0.50	8 : 0	4.50
	Total	73	44.72	-	-	2.88 : 0.60	69 : 4	4.41
Coolblue.nl	HD8821	176	-	-	-	2.65 : 1.31	-	4.41
	HD8824	109	-	-	-	2.86 : 1.33	-	4.35
	HD8834	55	-	-	-	3.00 : 1.05	-	4.49
	Total	340	-	-	-	2.77 : 1.27	-	4.41
Mediamarkt.nl	HD8821	56	-	-	-	2.43 : 1.21	51 : 5	4.18
	HD8824	16	-	-	-	3.19 : 0.94	15 : 1	4.50
	HD8834	13	-	-	-	2.54 : 0.31	11 : 2	4.54
	Total	85	-	-	-	2.59 : 1.02	77 : 8	4.29
Kieskeurig.nl	HD8821	21	-	-	1.51	3.24 : 2.05	12 : 9	3.62
	HD8824	5	-	-	0.75	3.60 : 1.20	2 : 3	4.00
	HD8834	9	-	-	0.25	3.22 : 0.89	8 : 1	4.67
	Total	35	-	-	1.09	3.29 : 1.63	22 : 13	3.94
Total*	**All**	565	44.85	2 : 1 (66 : 33%)	1.38	2.78 : 1.18	191:34 (85 : 15%)	3.96

*** total, averaged based on all available data**

Appendix C: Attributes identified by the researcher and validator

Identified by validator	Identified by researcher
Attributes validator	**Attribute researcher (matched to validator's)**
Desired temp.	Coffee temperature (adjustability)
Design	Design (aesthetics) Design look & feel (colour; materials)
Desired output	Cappuccino quality / taste Classic coffee option Coffee quality / taste Drink variety Espresso quality / taste Freshness of coffee Latte quality / taste Milk froth quality Warm water (tea) option / quality
Speed	Speed (making coffee) Speed (warming up) Speed (in general / undefined)
Sound	Grinders (sound) Milk frother sound Sound / noise level (in general)
Sustainable	Energy use machine
Ease of use User-friendliness	EOU Operation (ease of)
Whole beans option	Beans compatibility
Milk function	Milk frother EOU Milk frother performance/quality Milk frother spatters
Water reservoir handle	(Handle) water container (quality)
Filling water reservoir Water reservoir	Water container capacity (frequency of refilling) Water container ease of use/filling
Dispensing spout	(Adjustable) height coffee spout
Drip tray	Coffee residue bin capacity (f.o.e.) / wetness Drip tray capacity (frequency of emptying) Drip tray EOM (emptying)
Many Options	Number of options / settings Personalisation / adjustability
Instructions	Communication / notifications (through display)
Bean reservoir indicator	Sensor performance
Filling Bean reservoir	Bean container capacity (frequency of refilling) Bean reservoir ease of use/filling
Bean reservoir	Bean container quality (stream, lid)
Standby ON/Of button Time ON	Stand-by mode / on/off-button (light)
Installation	Installation (ease of)
Buttons	Buttons (comfort/EOU)
Big / small coffee option	Coffee volume (adjustability)

Perceived as external attributes (by validator)
Beans (choice. freshness)
Service (Philips; spare parts, reparation)
Service (delivery / retailer)

No occurrence in subset (validator)
Accessibility of parts
Cable (length)
Cappuccino not one step
Dishwasher proof parts
Pressure espresso
Ringing / bumping / shifting cups
Warm water point (take out / put in)

Display	Display (quality)	
Adjustable brew grinder	Coffee strength / grind (adjustability)	
Maintenance	EOM Frequency of maintenance Cost of maintenance / cleaning	
user manual (NL)	User manual	
Cleaning milk option	Milk frother cleaning (frequency) Milk frother cleaning (EO)	
Brewgroup replacement	Replacing brew unit	
Often cleaning	Frequency of cleaning	
Two cups at same time	2 cups option	
Manufacturer	Brand	
Descaling	Frequency of descaling	
Size	Size machine (compactness)	
Dripping / rinsing	Need for drip cup (for rinsing) Spatters	
Water waste	Leakage / Rinsing	
Quality machine	Defects Performance machine (function) Quality machine Value for money	
Price (expectations)	Price machine	
Quality of functions and parts	Brew group quality Quality of containers	
Saving settings	Saving options (profiles)	
Settings	Settings (ease of)	
Milk option	Milk solution (separate / easy)	
Option for powder coffee	Ground coffee (decaf) option Ground coffee EOU (small opening)	
Cleaning	Brew unit cleaning EOC (automatically) Speed of cleaning	
Total attr. validator: 45	**Total attr. researcher: 84**	

Decisions made in accordance with validator:

Continued attributes (coloured green in list): 82

** Most detailed attribute labels of matched list (coloured green)*

** Attr. not occured in validation subset*

Excluded attributes (coloured red in list): 3

** External attributes*

Appendix D: The number of reviews required for validation

As described in paragraph 3.3.1, validation of the attribute identification step was performed based on a subset containing 85 out of 565 randomly selected reviews (~15%) from the dataset. The formula of Nielsen & Landauer (1993) provides a widely-applied, suitable method to determine the number of users needed in user tests (Sauro & Lewis, 2016). The table below shows how the formula is applied to this validation test and how values are calculated.

Concept (Nielsen & Landauer, 1993)	**Concept (this study)**	**Symbol**	**Formula**	**Value**
Average number of issues (usability problems) addressed by a single user	Average number of attributes (pros and cons) mentioned by a single reviewer	x	total no. of attributes identified / total no. of reviews	3,96
Total issues in the design/product	Total attributes (issues) of the products, identified by researcher	N		84
No. users in the usability test	No. of reviews in the validation test	n		85
% of all issues, discovered while testing a single user	% of all attributes, identified in one online product review	L	x / N	4,71%
No. issues found by this test	No. attributes identified by this test	Y	$N*(1-(1- L)^n)$	82,61
% of all issues, discovered by this test	% of all attributes, identified by this test	C	Y / N	98,35%

In the context of this study, the number of reviews represent the number of users that 'tested', i.e., used, the product, and the number of issues detected is the number of pros and cons mentioned in the review. It was established that 98.35% (82.61 attributes) of all identified product attributes from the complete dataset (84) is to be covered by 85 consumer reviews. This calculation is based on the average number of product attributes addressed by a single consumer (3.96).

Appendix E: Instructions for validation of attribute identification

Goal and definition

The goal of this validation task is to identify the topics that consumers mention in their product reviews.

The topics mentioned by consumers are referred to as product attributes, i.e., subjective characteristics of a product, that are perceived and evaluated by consumers, expressed in their product evaluations arising from usage situations (Abraham-Murali & Littrell, 1995; Thompson et al., 2005; Kangale et al., 2016). Product attributes can be perceived as benefits or costs, which consumers express in their pro and con sections of reviews. Examples of product attributes are perceived aesthetics, performance, personalisation, and ease-of-use. It is chosen to identify product attributes from consumer reviews to be able to assess the product from the consumer's perspective.

Dataset

In the Excel sheet attached, you will find a subset of my data, consisting of 85 randomly selected online product reviews, thus, including an unknown number of reviews for the three different fully automatic espresso machines of Philips that were studied. Each row represents a unique review, whereas columns are the characteristics of a review. The subset only includes those characteristics that are needed to perform the validation task, i.e., review number, review title, free text area, pro section, con section.

Task

I would like you to transform every pro and con mentioned into product attributes. Previous literature has shown that pro and con sections generally provide a summary of the more elaborate product evaluations that the reviewer has written in the free text area. Therefore, only the pros and cons of each review need to be transformed into product attributes. However, when you find a product attribute poorly described in the pro/con section, you are free to read the free text area and title of the review to be able to define an appropriate attribute name. The attribute identification should be completely based on your own interpretation, based on your own usage experiences with your Philips full automatic espresso machine.

30 empty attribute columns are inserted in the header row, which you will rename to represent the unique product attributes that you identified from the reviews. The number 30 is randomly chosen, and does not imply that you will have to identify exactly 30 attributes. In fact, you might identify way less or more attributes. Please feel free to insert additional attribute columns if needed.

The steps to perform review-wise (row by row), for each phrase written in the pro/con sections of a review, are:

- For each product attribute you perceive as conceptually distinct, rename a column by replacing "[Attribute name]" with the name you think best represents the product attribute.
- For each product attribute you perceive as conceptually similar to a previously named product attribute, no new attribute name needs to be added.

Thank you very much! Please feel free to contact me in case of any unclarities or questions.

Appendix F: Instructions for validation of attribute categorisation

Thank you for filling in this form and helping me graduate from the master's Innovation Management (TU/e). Please note that, to participate, you should be owner or frequent user of a full automatic espresso machine of the brand Philips.

My master thesis aims to improve Philips' full automatic espresso machines with the objective of capturing a bigger share of the coffee machine market in the Netherlands. Consumers' online product reviews were analysed to detect specific consumer needs that should be addressed in Philips' new product development process.

One step in my analysis is to classify different attributes (characteristics) of full automatic espresso machines into three categories: capability, usability and maintainability. Product attributes were extracted from online product reviews, i.e., consumers' product evaluations arising from using the product in their specific contexts. I would like to ask you to classify the product attributes into categories based on your own usage experiences. This will take up to 10 minutes of your time and will be incredibly helpful to my graduation process.

Please read the following definitions of the attribute categories carefully before answering the questions.

Capability (=vermogen/kwaliteit/bekwaamheid)

- The product’s ability to perform the desired functions
- Examples of capability attributes: speed of making coffee, display quality

Usability (=bruikbaarheid/gebruiksvriendelijkheid)

- The ease of learning and using the desired functions
- Examples of usability attributes: ease of operation, milk frother ease of use

Maintainability (=onderhoudbaarheid/onderhoudsvriendelijkheid)

- The ease of cleaning and maintaining the product to keep the desired functions in working condition
- Examples of maintainability attributes: frequency of maintenance, descaling

There are no right or wrong answers. The categorisation should be based upon your personal usage experiences and perceptions, as I would like to compare different points of view of multiple consumers. Only when an attribute does not belong to any of these categories, or would rather belong to two categories, choose the "none / multiple" option and specify your thoughts in the text area at the end.

In case of any unclarities or questions, feel free to contact me. Thank you very much!

Appendix G: Results of attribute categorisation validation study

Respondent characteristics						
Respondent number	1	2	3	4	5	6
Role	Consumer	Consumer	Consumer	Consumer	Consumer	Reseacher
Timestamp	9-12-2017 15:18	12-12-2017 12:59	13-12-2017 14:02	13-12-2017 14:04	17-12-2017 22:31	15-12-2017 08:35
Sex	Male	Male	Male	Male	Female	Female
Age	32	21	25	21	57	27
Location	Eindhoven	Eindhoven	Eindhoven	Eindhoven	Veldhoven	Eindhoven
Education	High	High	Mid	High	Mid	High
Product type	HD8821/01	EP3550/00	HD8925/01	HD8824/01	HD8847/01	None

\

Attribute categorisation into capability, usability and maintainability								
Attributes	**1**	**2**	**3**	**4**	**5**	**6**	**F**	**#**
(Adjustable) height coffee spout	C	C	x	C	U	C	C	3
(Handle) water container (quality)	C	C	C	U	C	C	C	4
2 cups option	C	C	C	U	C	U	C	4
Accessibility parts	C	M	U	U	M	M	M	2
Bean container capacity (frequency of refilling)	U	U	M	U	C	U	U	3
Bean container ease of use/filling	U	U	U	M	U	U	U	4
Bean container quality (stream, lid)	C	C	C	C	C	C	C	5
Beans compatibility	C	C	C	C	C	C	C	5
Brew group quality	C	C	M	C	C	C	C	5
Brew unit cleaning	M	M	M	M	M	M	M	5
Buttons (comfort/ease of use)	U	U	C	U	U	U	U	4
Cable (length)	C	C	x	U	U	C	U	2
Cappuccino not available in one push on button	U	U	C	U	C	C	U	3
Cappuccino quality / taste	C	C	x	C	x	C	C	3
Classic coffee option	C	C	U	C	C	C	C	4
Coffee quality / taste	C	C	C	C	x	C	C	4
Coffee residue bin capacity (emptying frequency)	U	U	M	U	M	U	U	3
Coffee strength / grind (adjustability)	C	C	C	C	C	C	C	5
Coffee temperature (adjustability)	C	C	C	C	C	C	C	5
Coffee volume (adjustability)	C	C	C	C	C	C	C	5
Communication / notifications (through display)	U	U	U	C	U	U	U	4
Cost of maintenance / cleaning	M	M	x	M	M	M	M	4
Defects	C	M	C	C	C	C	C	4
Design (aesthetics)	C	C	C	C	x	C	C	4
Design look & feel (colour; materials)	C	C	C	C	U	C	C	4
Dishwasher proof parts	M	U	U	M	U	U	U	3
Display (quality)	C	C	C	U	C	C	C	4
Drink variety	C	C	C	C	U	C	C	4
Drip tray capacity (frequency of emptying)	M	M	x	U	C	U	U	2
Drip tray ease of use (emptying)	U	M	U	U	M	U	U	3
Ease of descaling	M	M	M	M	M	M	M	5
Energy use machine	U	C	x	C	C	C	C	3
Ease of cleaning	U	U	U	M	M	M	U	3
Ease of maintenance	M	M	U	M	M	M	M	4
Ease of use	U	C	C	U	U	U	U	3
Espresso quality / taste	C	C	C	C	x	C	C	4

*

Frequency of cleaning	M	M	U	M	M	M	M	4
Frequency of maintenance	M	M	M	M	M	M	M	5
Freshness of coffee	C	C	C	C	x	C	C	4
Grinders (sound)	C	C	x	C	M	C	C	3
Ground coffee (decaf) option	C	C	C	C	U	C	C	4
Ground coffee option ease of use (small opening)	U	U	U	U	C	U	U	4
Installation (ease of)	U	M	U	M	U	U	U	3
Latte quality / taste	C	C	C	C	C	C	C	5
Leakage / Rinsing	M	M	U	C	M	C	M	3
Milk froth quality	C	C	C	C	C	C	C	5
Milk frother cleaning (ease of)	M	M	M	M	M	M	M	5
Milk frother cleaning (frequency)	M	M	M	M	M	M	M	5
Milk frother ease of use	U	U	U	M	U	U	U	4
Milk frother performance/quality	C	C	C	C	C	C	C	5
Milk frother sound	C	C	U	C	C	C	C	4
Milk solution / option	C	C	U	U	C	C	C	3
Need for drip cup (for rinsing)	M	C	C	C	x	C	C	3
Number of options / settings	C	C	C	U	U	C	C	3
Operation (ease of)	U	U	U	C	M	U	U	3
Performance machine (function)	C	C	C	C	U	C	C	4
Personalisation / adjustability	U	C	C	U	U	C	U	3
Pressure espresso	C	C	C	C	C	C	C	5
Price machine	C	C	C	C	C	C	C	5
Quality machine	C	C	x	C	x	C	C	5
Quality of containers	C	C	C	C	C	C	C	5
Replacing brew unit	M	M	M	M	U	M	M	4
Ringing / bumping / shifting cups	U	C	C	C	U	C	C	3
Saving options (profiles)	U	C	U	U	C	C	U	4
Sensor performance	C	M	U	U	U	C	U	3
Settings (ease of)	U	U	C	C	U	U	U	3
Size machine (compactness)	U	x	C	U	x	C	C	2
Sound / noise level (in general)	C	C	x	C	x	C	C	3
Spatters in general (coffee/undefined)	C	C	M	U	C	C	C	3
Spatters milk frother	C	C	M	C	M	C	C	3
Speed (in general / undefined)	U	C	C	C	C	C	C	4
Speed (making coffee)	U	C	U	C	C	C	C	3
Speed (warming up)	C	C	U	C	C	U	C	4
Speed of cleaning	M	M	M	M	M	M	M	5
Stand-by light / button	U	x	C	C	C	C	C	3
Time untill stand-by modus	U	U	x	U	U	C	U	4
User manual	U	M	U	M	U	U	U	3
Value for money	C	C	C	C	x	C	C	4
Warm water (tea) option / quality	C	C	C	C	U	C	C	4
Warm water point (take out / put in)	C	C	C	U	C	C	C	4
Water container capacity (frequency of refilling)	U	M	M	U	U	U	U	3
Water container ease of use/filling	U	M	U	M	U	U	U	3

* 1-6 Respondent number

F Final categorisation decision

Number of respondents that categorised attribute in accordance with F

Appendix H: Section of sentiment analysis results in Excel

Review no.	Pros	Cons	Grinders (sound)	Sound / noise level	Spatters in general (coffee / undefined)	Speed (making coffee)	Speed (warming up)	2 cups option	Coffee temperature (adjustability)	Coffee strength / grind (adjustability)	Coffee volume (adjustability)	Coffee quality / taste	Cappuccino quality / taste	Espresso quality / taste	Fresh coffee	Beans compatibility	Ground coffee (decaf) option	Warm water (tea) option / quality	Drink variety	Milk froth quality	Milk solution/option	Milk frother performance / quality
1	Goede kwaliteit Goede prijs/kwaliteit Eenvoudig in gebruik Past perfect																					
2	Goede kwaliteit Eenvoudig in gebruik																					
3	Goede prijs/kwaliteit gebruikshandleiding vwb technisch onderhoud matig	handleiding vwb technisch onderhoud is matig bonenreservoirindicator werkt niet																				
4	Goede kwaliteit Goede prijs/kwaliteit Eenvoudig in gebruik																					
5	Goede kwaliteit																					
6	Goede kwaliteit Goede prijs/kwaliteit Eenvoudig in gebruik Past perfect																					
7	Goede prijs/kwaliteit Eenvoudig in gebruik																					
8	Goede kwaliteit Eenvoudig in gebruik verschillende koffie/specialiteiten	Hoge prijs handleiding niet in het nederlands																	1			
9	latte geweldig[SEP]vast melkkan[SEP]sterkte per kop regelbaar[SEP]melkkan goed zuiver te maken	bonenreservaat te klein								1											1	
10	Snel een lekkere kop koffie[SEP]In te stellen voor elke grote van kop[SEP]Prijs/kwaliteit [SEP]Mooi design[SEP]Snel opgewarmd	Koffiebonen moeten snel bijgevuld worden				1	1					1										

Review no.	Pros	Cons	Milk frother sound	Spatters milk frother	Size machine (compactness)	Design (aestethics)	Design look & feel (color; materials)	Quality machine	Value for money	Performance machine (function)	Defects	(Handle) water container (quality)	Bean container quality (stream, lid)	Display (quality)	Sensor performance	Stand-by mode / on/off-button (light)	Price machine	Personalisation / adjustability / profiles	Number of options / settings	(Adjustable) height coffee mouh	**Capability**
1	Goede kwaliteit Goede prijs/kwaliteit Eenvoudig in gebruik Past perfect				1			1	1												3
2	Goede kwaliteit Eenvoudig in gebruik							1													1
3	Goede prijs/kwaliteit gebruikshandleiding vwb technisch onderhoud matig	handleiding vwb technisch onderhoud is matig bonenreservoirindicator werkt niet							1						-1						
4	Goede kwaliteit Goede prijs/kwaliteit Eenvoudig in gebruik							1	1												2
5	Goede kwaliteit							1													1
6	Goede kwaliteit Goede prijs/kwaliteit Eenvoudig in gebruik Past perfect				1			1	1												3
7	Goede prijs/kwaliteit Eenvoudig in gebruik								1												1
8	Goede kwaliteit Eenvoudig in gebruik verschillende koffie/specialiteiten	Hoge prijs handleiding niet in het nederlands						1									-1				1
9	latte geweldig vast melkkan sterkte per kop regelbaar melkkan goed zuiver te maken	bonenreservaat te klein																	1		3
10	Snel een lekkere kop koffie In te stellen voor elke grote van kop Prijs/kwaliteit Mooi design Snel opgewarmd	Koffiebonen moeten snel bijgevuld worden			1			1	1												3

Review no.	Pros	Cons	Installation (ease of)	Settings (ease of)	User manual	Ease of use (in general)	Operation (ease of)	Communication/notifications (via	Buttons (comfort/EOU)	Milk frother EOU	Drip tray EOU (emptying)	Drip tray capacity (frequency of	Coffee residue bin capacity (emptying)	Leakage / Rinsing	Water container capacity (freq. Of	Bean container capacity (freq. of	Water container EOU	Bean container EOU	**Usability**
1	Goede kwaliteit Goede prijs/kwaliteit Eenvoudig in gebruik Past perfect					1													1
2	Goede kwaliteit Eenvoudig in gebruik					1													1
3	Goede prijs/kwaliteit gebruikshandleiding vwb technisch onderhoud matig	handleiding vwb technisch onderhoud is matig bonenreservoirindicator werkt niet			-1														-1
4	Goede kwaliteit Goede prijs/kwaliteit Eenvoudig in gebruik					1													1
5	Goede kwaliteit																		
6	Goede kwaliteit Goede prijs/kwaliteit Eenvoudig in gebruik Past perfect					1													1
7	Goede prijs/kwaliteit Eenvoudig in gebruik					1													1
8	Goede kwaliteit Eenvoudig in gebruik verschillende koffie/specialiteiten	Hoge prijs handleiding niet in het nederlands			-1	1													
9	latte geweldig[SEP]vast melkkan[SEP]sterkte per kop regelbaar[SEP]melkkan goed zuiver te maken	bonenreservaat te klein														-1			-1
10	Snel een lekkere kop koffie[SEP]In te stellen voor elke grote van kop[SEP]Prijs/kwaliteit[SEP]Mooi design[SEP]Snel opgewarmd	Koffiebonen moeten snel bijgevuld worden														-1			-1

Review no.	Pros	Cons	Ease of cleaning (automatical)	Frequency of cleaning	Milk frother cleaning (ease, frequency)	Accessability parts	Ease of maintenance (in general)	Frequency of maintenance	Descaling (ease, frequency)	**Maintainability**
1	Goede kwaliteit Goede prijs/kwaliteit Eenvoudig in gebruik Past perfect									
2	Goede kwaliteit Eenvoudig in gebruik									
3	Goede prijs/kwaliteit gebruikshandleiding vwb technisch onderhoud matig	handleiding vwb technisch onderhoud is matig bonenreservoirindicator werkt niet								
4	Goede kwaliteit Goede prijs/kwaliteit Eenvoudig in gebruik									
5	Goede kwaliteit									
6	Goede kwaliteit Goede prijs/kwaliteit Eenvoudig in gebruik Past perfect									
7	Goede prijs/kwaliteit Eenvoudig in gebruik									
8	Goede kwaliteit Eenvoudig in gebruik verschillende koffie/specialiteiten	Hoge prijs handleiding niet in het nederlands								
9	latte geweldig[SEP]vast melkkan[SEP]sterkte per kop regelbaar[SEP]melkkan goed zuiver te maken	bonenreservaat te klein			1					1
1	Snel een lekkere kop koffie[SEP]In te stellen voor elke grote van kop[SEP]Prijs/kwaliteit [SEP]Mooi design[SEP]Snel opgewarmd	Koffiebonen moeten snel bijgevuld worden								

Appendix I: Instructions for sentiment analysis validation

Goal and definition

The goal of this validation task is to score the opinions that consumers express in their reviews towards the previously identified product attributes.

A consumer's opinion towards a product attribute will be represented by a positive (1), negative (-1), or neutral ( ) score, based on whether the attribute was mentioned in the review's pro section, con section, or was not mentioned in neither of both sections.

Dataset

In the Excel sheet attached, you will find a subset of my data, consisting of 85 randomly selected online product reviews, thus, including an unknown number of reviews for the three different fully automatic espresso machines of Philips that were studied. Each row represents a unique review, whereas columns are the characteristics of a review. The subset only includes those characteristics that are needed to perform the validation task, i.e., review number, pro section, con section, and a list (range of columns) of 61 previously identified product attributes. The product attributes are divided into three underlying categories of consumer needs: capability (red), usability (yellow/orange), and maintainability (green).

Task

I would like you to transform every pro and con mentioned into product attribute scores. Pros are the aspects of the product that the reviewer indicated to perceive as positive, cons are the negatively perceived aspects.

61 attribute columns were inserted in the header row, which you will use to fill in the attributes scores obtained from each review (row).

The step to perform review-wise (row by row), for each phrase written in the pro/con sections of a review, is:

- Assign a score to each product attribute (pro/con phrase), by entering "1" or "-1" in the cell that connects the review (row) with the attribute (column) in question:

  1 Positive (mentioned in pro section)
  -1 Negative (mentioned in co section)

Thank you very much! Please feel free to contact me in case of any unclarities or questions.

Appendix J: Results of sentiment analysis validation

			Researcher			Validator			Difference		
Review no.	**Pros**	**Cons**	**Capability**	**Usability**	**Maintainability**	**Capability**	**Usability**	**Maintainability**	**Capability**	**Usability**	**Maintainability**
3	Goede prijs/kwaliteit gebruikshandleiding vwb technisch onderhoud matig	handleiding vwb technisch onderhoud is matig bonenreservoirindicator werkt niet	0	-1	0	0	-1	-1			x
6	Goede kwaliteit Goede prijs/kwaliteit Eenvoudig in gebruik Past perfect		3	1	0	2	1	0	x		
10	Snel een lekkere kop koffie[SEP]In te stellen voor elke grote van kop[SEP]Prijs/kwaliteit[SEP]Mooi design[SEP]Snel opgewarmd	Koffiebonen moeten snel bijgevuld worden	7	-1	0	6	-1	0	x		
12	gemakkelijke bediening[SEP]stil[SEP]lekkere koffie		2	1	0	2	1	0			
15	Gemakkelijk in gebruik[SEP]Mooi design[SEP]Lekkere cappuccino	Wel wat veel lawaai, vooral bij melk opschuimen	1	1	0	1	1	0			
16	Goede kwaliteit prijs verhouding Geluid van malen bonen valt erg mee Gebruiksvriendelijk Memory + afstelling	Net iets te klein waterreservoir	3	1	0	3	0	0		x	
18	Eenvoudig voor gebruik[SEP]Koffie smaakt goed[SEP]Compact	Behuizing kon van betere materiaal gemaakt worden[SEP]Vaak schoonmaken en vaak onderhoud[SEP]Kleine reservoirs voor water en koffie	1	-1	-2	1	-1	-2			
25	Lekkere cappuccino[SEP]Lekkere koffie		2	0	0	2	0	0			
29	Snel een bakje koffie[SEP]Goede prijs	Heeft veel onderhoud nodig	2	0	-1	2	0	-1			
30	compact eenvoudig te bedienen veel mogelijkheden gezien de prijs	compact, dus niet zoveel water/bonen voorraad geen ProfielOpslag, bij wel PO geen gemalen koffie bonencompartiment stroomt soms niet goed door	1	-2	0	1	-2	0			

34	Eenvoudig in gebruik Goed met afvalscheiding koffiedrab en water en schoonhouden Prima koffie en temperatuur	Maakt veel lawaai Klein reservoir voor bonen	1	2	1	1	2	1			
36	Mijn kinderen kunnen ook chocolademelk opschuimen Makkelijk te bedienen		2	1	0	2	1	0			
45	Verse koffie Snelklaar Handig melkvakje		3	0	0	3	0	0			
56	Smaak instelbaar Makkelijk te reinigen Cappuccino schuim perfect Alle onderdelen melkkan uit elkaar te halen: optimaal schoonmaken	Melk maximaal een paar dagen in kan houdbaar daarna echt schoonmaken. Maar dat is met een kan van een ander merk ook zo	3	0	1	3	0	1			
62	Maakt heerlijke Capuchino Gemakkelijke bediening. Mooie vormgeving. Neemt niet zoveel plaats in.	Meer onderhoud dan gedacht.	3	1	-1	3	1	-1			
74	Makkelijke bediening		0	1	0	0	1	0			
78	Heerlijke koffie ! , Via internet is de volledige handleiding te downlo	Nog niet ontdekt !	1	1	0	1	1	0			
84	Goed warm , Lekker van smaak , Makkelijk schoon te maken , Goede melknopschuimer	Geen	3	0	1	3	0	1			
85	Zeer gebruiksvriendelijk.	Maar 1 kopje koffie per keer.	-1	1	0	-1	1	0			
90	makkelijk in gebruik, gebruik van verse bonen		1	0	0	1	1	0		x	
102	Niet zo'n heel duur apparaat	Malen van de bonen maakt heel veel herrie De kopjes rinkelen op het plateautje Stoompijpje en slangetje spetteren	-1	0	0	-1	0	0			
112	Makkelijk in gebruik , Hete koffie , Goede melk opschuimer		2	1	0	2	1	0			
123	lekkere koffie	Matige kwaliteit veel lawaai traag (lang wachten voor koffie kleine waterreservoir (5 kopjes)	-2	-1	0	-3	-1	0	x		
132	Goede kwaliteit Goede prijs/kwaliteit Voordelige koop Eenvoudig in gebruik Past perfect		4	1	0	4	1	0			
135	Goede kwaliteit Goede prijs/kwaliteit Eenvoudig in gebruik	maakt meer lawaai omdat het de bonen vermaalt	1	1	0	0	1	0	x		
152	Waterreservoir makkelijk Goede koffie	Gaat niet automatisch uit na bepaalde tijd	0	2	0	0	2	0			

158	Lekkere koffie en espresso Prima prijs / kwaliteit Makkelijk in onderhoud Makkelijke bediening	Beperkte capaciteit koffiebonenmagazijn	3	0	1	3	0	1			
165	opschuimer werkt prima Goed instelbare smaak Compact	ziet er iets goedkoop uit	2	0	0	2	0	0			
187	Is snel opgewarmd makkelijk in te stellen Hete koffie		2	1	0	2	1	0			
191	smaak	veel verbruik	1	-2	0	1	-1	0		x	
194	Compact toestel Eenvoudig in gebruik Groot koffiebonen en water reservoir	Maakt tamelijk veel lawaai Vraagt geregeld onderhoud	0	3	-1	0	2	-1		x	
200	Gemak en eenvoud	Wel wat groot	-1	1	0	-1	1	0			
211	Makkelijk in gebruik Goed schoon te houden		0	1	1	0	1	1			
218	Gemakkelijk in gebruik Heerlijke capucino Watervolume aanpasbaar Reinigen en water bijvullen is kinderspel	Tempertuur water mag beter	1	2	1	1	2	1			
221	Lekker bakkie Makkelijk in gebruik Compact design	**Kopje wordt nog niet automatisch aan tafel geserveerd**	3	1	0	3	1	0			
224	Degelijk apparaat, makkelijk te bedienen Beste prijs kwaliteit verhouding Goede heerlijke koffie! Makkelijk te reinigen Via Coolblue snel in huis!	Melk opschuimen maakt veel geluid (maar wen je aan) Lekbak snel vol Melkslang vaak reinigen om aanslag te voorkomen Temperatuur had wat heter gemogen	2	0	-1	1	0	-1	x		
227	Lekkere cappuccino Makkelijk te bedienen Sterkte koffie instellen		2	1	0	2	1	0			
247	Lekkere koffie op maat Gebruiksvriendelijk		2	1	0	2	1	0			
250	Makkelijk schoon te houden. Gebruiksvriendelijk.		0	1	1	0	1	1			
263	lekkere koffie	gekookte melk ipv opgeschuimde	-1	0	0	-1	0	0			
264	makkelijk ingebruik te nemen, snel een kopje koffie of espresso	lekt best veel water, koffie niet heel warm	0	-1	0	0	-1	0			
268	eenvoudig in gebruik, extra klepje om de bonen vers te houden, goede melkopschuimer, mooi apparaat in de keuken	instellen grofheid bonen	2	0	0	2	0	0			
281	Prijs Lekkere koffie Apparaat is relatief klein en ziet er mooi uit Display is duidelijk	Lekbakje redelijk snel vol Iets wat lawaaierig bij het malen	4	0	0	3	0	0	x		

285	Mooi design Goed heet tapwater en stoom Makkelijk te reinigen Voldoende grote watertank		2	1	1	2	1	1			
294	melkopschuimer smaak van de koffie	Knipperlichtje	1	0	0	1	0	0			
295	Goede prijs Beste koop consumentenbond		2	0	0	2	0	0			
315	makkelijk te bedienen	lawaaierig	-1	1	0	-1	1	0			
332	Eenvoudig in gebruik , Lekkere koffie		1	1	0	1	1	0			
337	Melkopschuimer, makkelijk te bedienen	fijnere maling zou de smaak ten goede komen	-1	1	0	-1	1	0			
351	Bescheiden formaat , Prima espressokoffie , Sterkte makkelijk aanpasbaar , Snel voor 1 kopje koffie	Lawaaiig	3	0	0	3	0	0			
358	Goede kwaliteit Goede prijs/kwaliteit Voordelige koop	Kwetsbaar maakt relatief veel lawaai.	1	0	0	0	0	0	x		
364	Goede kwaliteit Goede prijs/kwaliteit Eenvoudig in gebruik	lekbak snel vol	2	0	0	2	0	0			
367	Goede kwaliteit Goede prijs/kwaliteit Voordelige koop Eenvoudig in gebruik snel in gebruik	best veel lawaai van de maler bij opwarmen komt er een klein laagje water vooraf	3	1	0	3	1	0			
370	Goede kwaliteit	Lastig in gebruik	1	-1	0	1	-1	0			
376	Goede kwaliteit Goede prijs/kwaliteit Voordelige koop	Lastig in gebruik	3	-1	0	3	-1	0			
390	Prijs-kwaliteit Warme koffie Stil	Klein koffiebonenreservoir	3	-1	0	3	-1	0			
398	Handig in gebruik Mooi design Super lekkere koffie natuurlijk	Hij maakt best wat geluid	1	1	0	1	1	0			
399	weinig lawaai reinigen apparaat is erg eenvoudig Eenvoudig in te stellen eenvoudige bediening koffie smaakt goed	bonen lastig bijvullen, door vaste deksel Komt niet heel solide over hengsel waterbak licht uitgevoerd	-2	1	1	0	1	1	x		
412	Goede koffie Gemakkelijk bonen bijvullen Mooi design	 Cappucino zetten is onmogelijk, melk spuit overal heen Veel geluid	-1	0	0	-1	0	0			
413	Goeie koffie Maakt niet al teveel geluid Klein Kan alles!	Nu komt iedereen bij mij koffie drinken	4	0	0	4	0	0			
418	de smaak voor de prijs makkelijk in gebruik dosering zelfs mijn reuzekoffiemok past er onder	beetje groot uitgevallen vaak koffiegruis legen vrij onhandig bonenvullen	4	-1	0	4	0	0		x	

426	Makkelijk in gebruik Snel een lekkere kop koffie	Bij twee kopjes tegelijkertijd zit in het ene kopje minder dan in het andere Best lawaaiig	-1	1	0	-1	1	0			
436	Super duidelijk Temperatuur super Goed in te stellen maalgraad		2	1	0	2	1	0			
438	Thee zetten met hetzelfde apparaat zonder koffiesmaak! Makkelijk te bedienen. Makkelijk te reinigen. Smal apparaat. Best getest consumentenbond 216.	Melkschuimer aan de voor ons verkeerde kant, jammer dat hier geen keuze in is.	3	1	1	3	1	1			
444	Makkelijk in gebruik Mooi/strak Design		1	1	0	1	1	0			
449	Maakt niet veel herrie als hij bonen maalt Gebruiksvriendelijk Staat mooi te stralen op m'n aanrecht	Ik kan geen minpunten bedenken	2	1	0	2	1	0			
450	Apparaat is niet te groot Lekkere koffie Gemakkelijk schoon te maken	Opstarten 1e x was lastig en duurde lang Melk schuimt niet maar kookt over	0	-2	1	0	-2	1			
455	Prijs Koffie smaakt lekker Eenvoudig te reinigen	Geen	2	0	1	2	0	1			
458	Een echt koffiezetapparaat Makkelijk in gebruik Geen bijzonder onderhoud nodig Lekkere koffie Ook heet water, en mogelijkheid om melk op te warmen	Geluid is wel aanwezig als hij iets doet.	3	1	1	3	1	1			
467	design gebruiksvriendelijk ruim reservoir		1	3	0	1	2	0		x	
470		Koffie niet echt heet	-1	0	0	-1	0	0			
485	Kleiner dan vergelijkbare apparaten Gebruiksgemak	Klein water reservoir Klein bonen reservoir	1	-1	0	1	-1	0			
501	heerlijke koffie! makkelijk aan te passen aan eigen voorkeur		2	0	0	2	0	0			
511	Heerlijke koffie	als je een klein koffietasje hebt spat het wel een beetje, de uitloop zou iets verder moeten kunnen uitrekken	-1	0	0	-1	0	0			
512	Gemakkelijk in gebruik Prima maat Lekkere koffie	Rode lampje blijft knipperen bij stand bij	1	1	0	1	1	0			

522	Espresso is lekker en makkelijk Snel klaar Simpele bediening	Gewone koffie zou sterker mogen, nog wat proberen met andere bonen en maling Geen Nederlandstalige handleiding bijgesloten, alleen een NL snelstart gids Rood lampje blijft knipperen en stand-by	-1	1	0	0	1	0	x			
524	Niet te groot Ziet er mooi uit Betaalbaar Instelbaar	Melkopschuimen ingewikkeld.	4	-1	0	3	-1	0	x			
530	Compact model, goede kwaliteit onderdelen. Geen plastic feel. Lekkere koffie	Bonenreservoir loopt niet helemaal leeg. Er blijven bonen liggen.	3	0	0	3	0	0				
533	Heerlijke koffie Makkelijk te bedienen Makkelijk schoon te houden	Waterreservoir snel leeg Zit geen waterfilter bij, wel los te krijgen	1	0	0	1	0	0				
536	makkelijk in gebruik	wat lawaai tijdens het malen van de koffie	-1	1	0	-1	1	0				
540	Makkelijk in gebruik Goede prijs	Gaat al na 15 minuten weer uit	0	1	0	0	1	0				
544	compact / smal design smaak snel prijs	geluid afwerking	2	0	0	2	0	0				
559	Zeer compact geluid is niet te hard/vervelend	Als het water op is en je kopje is nog niet vol dan gaat de water afgifte niet verder als het reservoir weer gevuld is. Dus moet je een nieuwe kopje zetten.	1	-1	0	1	-2	0		x		
562	Heerlijke koffie Gebruiksgemak	Nog niet tegengekomen	1	1	0	1	1	0				
563	snel koffie ook espresso zonder trucs uit te halen eenvoudig in gebruik		1	2	0	1	2	0				
									10	7	1	92,94%

Appendix K: Final list of product attributes, categorised

Capability	
(Adjustable) height coffee spout	Milk froth quality
(Handle) water container (quality)	Milk frother performance / quality
2 cups option	Milk frother sound
Bean container quality (stream, lid)	Milk solution/option
Beans compatibility	Number of options / settings
Cappuccino quality / taste	Performance machine (function)
Coffee quality / taste	Personalisation / adjustability / profiles
Coffee strength / grind (adjustability)	Price machine
Coffee temperature (adjustability)	Quality machine
Coffee volume (adjustability)	Sensor performance
Defects	Size machine (compactness)
Design (aesthetics)	Sound / noise level
Design look & feel (colour; materials)	Spatters in general (coffee / undefined)
Display (quality)	Spatters milk frother
Drink variety	Speed (making coffee)
Espresso quality / taste	Speed (warming up)
Fresh coffee	Stand-by mode / on/off-button (light)
Grinders (sound)	Value for money
Ground coffee (decaf) option	Warm water (tea) option / quality

Usability	**Maintainability**
Coffee residue bin capacity (emptying)	Accessibility parts
Drip tray EOU (emptying)	Descaling (ease, frequency)
Bean container capacity (freq. of refilling)	Ease of cleaning (automatically)
Communication/notifications (via display)	Ease of maintenance (in general)
Bean container EOU	Frequency of cleaning
Drip tray capacity (frequency of emptying)	Frequency of maintenance
Buttons (comfort/EOU)	Milk frother cleaning (ease, frequency)
Ease of use (in general)	
Installation (ease of)	
Leakage / Rinsing	
Milk frother EOU	
Operation (ease of)	
Settings (ease of)	
User manual	
Water container capacity (freq. of refilling)	
Water container EOU	

Appendix L: Significant attribute effects on satisfaction

Coefficients: effects of attributes on satisfaction for HD8821 (14 features)					
	Standardised Coefficients				
	Beta	t	Sig.	Mean	Std. dev.
Milk frother performance / quality	,194	3,430	,001	-,01	,189
EOU	,186	3,261	,001	-,04	,203
Size machine (compactness)	,165	2,929	,004	-,01	,114
Grinders (sound)	,141	2,519	,012	-,05	,440
Bean container EOU	,131	2,062	,040	-,02	,197
Sound / noise level	,123	2,171	,031	-,01	,099

Coefficients: effects of attributes on satisfaction for HD8824 (16 features)					
	Standardised Coefficients				
	Beta	t	Sig.	Mean	Std. dev.
Milk frother sound	,300	2,982	,004	-,02	,132
EOC (automatical)	,266	3,116	,002	-,01	,108
Milk frother performance / quality	,215	2,270	,025	-,04	,185

Coefficients: effects of attributes on satisfaction for HD8834 (21 features)					
	Standardised Coefficients				
	Beta	t	Sig.	Mean	Std. dev.
Design (estethics)	,412	2,796	,009	0,00	0,000
Settings (ease of)	,266	2,310	,028	,01	,185