Erasmus University Rotterdam

Erasmus School of Economics

Master in Marketing

- Master Thesis -

Design Aesthetics in Mobile Interface Design

-

The effect of design aesthetics on perceived usability and willingness to buy

Student:
Daniël Tolman
356886

Supervisor:
Muhammad Asim

Academic Year 2015-2016

Abstract

In the nineties, the conventional idea among interface designers was that no amount of 'sparkle' in the presentation of products could overcome a site design with poor navigation features. However, in the last decade this idea changed as numerous studies found evidence that a high perceived aesthetic appeal of a website is actually capable of overriding users' poor usability experience. By conducting an experiment, this study attempts to find the separate influence of eight aesthetical screen attributes on the perceived usability of a site or application and the subsequent users' willingness to buy. Results indicate that design aesthetics can indeed affect the perceived usability of a mobile retailer's application. Furthermore, evidence is found that users' willingness to buy is also influenced by both design aesthetics as well as the perceived usability of the application. Regarding the separate effects of the screen attributes, three attributes (screen density, color design, product rating) were found to have significant influence on both dependent variables. Interestingly though, the direction of these separate effects is rather different than the hypothesized effects, which leaves room for further research on this matter.

Table of Contents

1. Introduction

1.1 Context of research and problem definition

Mobile commerce or mobile marketing, also referred to as m-commerce, is a rapidly growing part of marketing, already accounting for a considerable part of global retail sales. Mobile marketing can be defined as the two-way or multi-way communicational and promotional activities between a firm and its customers, using a mobile device, medium or technology (Shankar & Balasubramanian, 2009). According to data company statista (2014), the number of worldwide smartphone users is forecast to reach over two billion in 2016. By 2019, it is expected that this number has already surpassed 2.6 billion users worldwide. Next to that, the number of tablet users is expected to reach 1.2 billion in 2016 and close to 1.5 billion in 2018 (eMarketer, 2015). These forecasts confirm that we are living in an era that becomes increasingly digital.

The rapidly growing use of mobile devices creates massive opportunities for retailers to operate in a third channel, next to the traditional brick and mortar- and e-commerce-channels. By 2020, U.S. m-commerce sales are expected to reach $252 billion, which will account for 49 percent of total online sales (Wu, 2015). A new channel and increasing technological possibilities also creates huge opportunities for mobile advertising. In 2011, Gartner already expected that mobile advertising revenue worldwide would increase 1167% from 2010-2015. It is evident that companies in the retail industry can no longer ignore the emerging m-commerce channel.

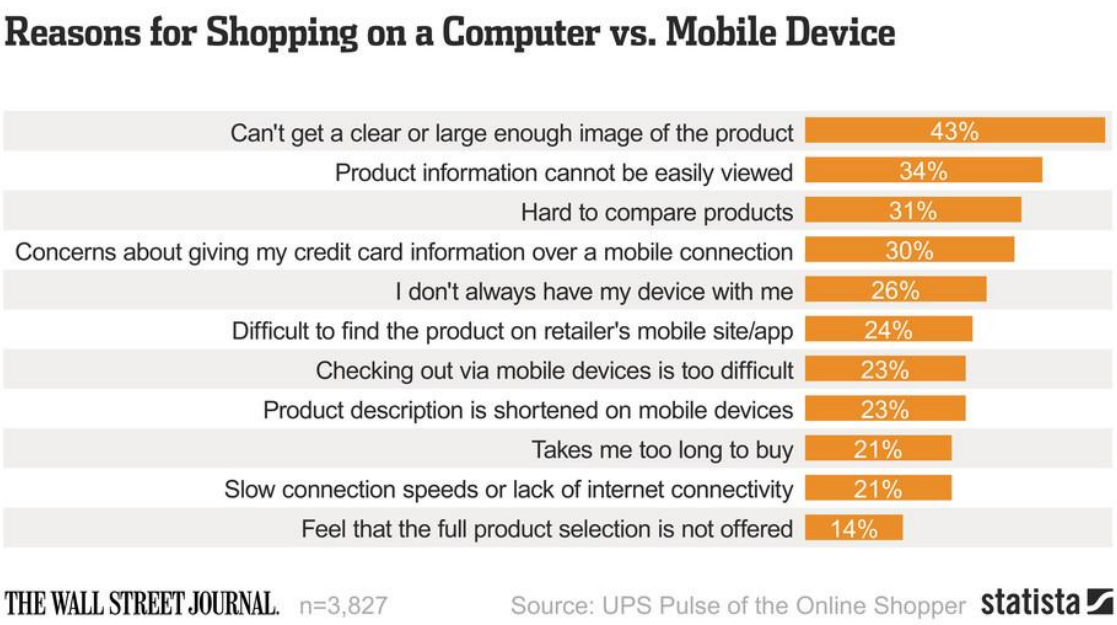
However, m-commerce is a relatively new channel with dynamics that are significantly different from those of other channels. Whereas e-commerce has been extensively researched in the past decades, there is still much research to be done in the field of m-commerce. A global research of Criteo (2015) found that western countries are having trouble to achieve acceptable conversion rates in the smartphone retail channel. In contrary, the conversion rates in Asian countries are thrice as high than the rates in their western counterparts. Next to this, the number of products viewed per user is significantly higher in Asian countries than in western countries (figure 1 appendix A). These findings raise several questions for mobile marketers. How can these significant differences be explained? Are Asian countries more likely to adapt new technologies or are there other factors that explain these findings?

Based on Hofstede's (1984, 1997) theory, who identified five cultural dimensions that could explain one's thinking, feelings and acting based on cultural values, difference in culture could explain at least a part of the abovementioned differences. Straub (1994) and Straub, Keil & Brenner (1997) suggest that cultural differences can indeed affect the adoption of new technologies and the attitude towards these technologies. Furthermore, a study by Markus and Gould (2000) found significant differences in user interface design across different cultures, suggesting that interface designs are perceived differently based on one's

cultural dimension. The findings of these studies suggest that difference in culture can influence the adoption of a new technology. However, it seems unlikely that this is the only factor that explains the fact that conversion rates in Asian countries are thrice as high than the rates in western countries.

Another explanation may lie in understanding the customer's needs in regard to m-commerce. Do Asian marketers have a better understanding of these needs and are they able to design their applications and mobile browsers more effectively according to these needs? Data company statista conducted a survey in the U.S. asking respondents for their opinion on shopping on a computer versus shopping on a mobile device. The findings provided some interesting insights in regard to the customer's adoption of mobile commerce (McCarthy, 2014):

Figure 1: Issues in mobile commerce adoption



These findings suggest that usability, safety, privacy and technological issues hinder people to adopt mobile commerce. Prior research in the field of m-commerce has also identified these various issues as possible obstacles for the adoption of this new technology (Tarasewich, Nickerson & Warkentin, 2002; Tarasewich, 2003a; Shankar & Balasubramanian, 2009; Shankar et al., 2010).

Even though Asian countries provide better figures, m-commerce in Asia also still has a long way to go to catch up with the conversion rates of e-commerce (Criteo, 2015). Therefore, it is critical to understand the above mentioned issues that currently hinder m-commerce from being adopted. What causes this issues and how can they be solved? The answer on this question is essential in order to understand the dynamics of mobile commerce and to enhance the adoption and effectiveness of it. Looking at the findings of statista (McCarthy, 2014), it can be noticed that many of the top rated issues are issues related to the usability

of an application or mobile browser. Considering the fact that smartphone screens are significantly smaller than desktop screens, it makes sense that the usability of smartphone screens is low compared to other, bigger devices. In the past, extensive research has been conducted to examine the usability of websites and the effectiveness of e-commerce (Baty & Lee, 1995; Lohse & Spiller, 1998a; Lohse & Spiller, 1998b, Shneidermann, 1998). Nowadays, the e-commerce market has matured and many years of interface design research has improved the effectiveness of websites, for a major part due to the high usability of the sites (statista, 2016). Can improved interface designs also lead to an increased adoption and effectiveness of mobile commerce? It definitely is an interesting question that needs further in-depth research.

One part of user interface design that has gained a lot of attention in the past decade are the design aesthetics of websites and applications. Design aesthetics can be defined as the balance, emotional appeal or aesthetic of a (mobile) website or application, which may be expressed through colors, shapes, font type, music or animation (Cyr et al., 2006). According to recent studies, design aesthetics can significantly influence perceived usability (PU), perceived ease-of-use (PEOU), loyalty, trust and other types of emotional reactions (Cyr et al., 2006; Hartmann, et al., 2007; 2008; Li & Yeh, 2010; Nanda et al., 2008, Sauer & Sonderegger, 2009; Sonderegger & Sauer, 2010). These findings may hold the key to increase the adoption rate and effectiveness of mobile commerce. Next to that, it is interesting to compare design aesthetics of Asian mobile websites and applications with their western counterparts. Are there any notable differences in design and can they explain the differences in conversion rates? This study will attempt to provide more insight in design aesthetics in the mobile commerce environment and the influence on customer's emotional reactions. The central research question of this thesis will be as follows:

Can design aesthetics influence the perceived usability of mobile websites and applications and can it subsequently increase the customer's willingness to buy?

1.2 Research objective and relevance

Designing a successful application or mobile browser or application is a process in which companies face many challenges. Consumers demand the same functionalities as they have on their desktops, but all these functionalities have to be merged into a considerably smaller screen. Moreover, this screen highly depends on ease of use, wireless connection, battery lifetime and portability (Shankar et al., 2010; Samet et al., 2011; Tarasewich, Nickerson & Warkentin, 2002; Tarasewich, 2003a). Next to these challenges, companies have to deal with customers who experience a higher perceived risk in privacy and security when using mobile devices (Webcredible, 2012).

Basically, mobile interface designers are faced with the enormous task to optimize the user interface on devices with physical limitations, while they can still ensure the customer's privacy and security. Next to that, it is expected that these mobile browsers and apps yields the same effectiveness as e-commerce activities. As can be concluded from the relatively low conversion rates in 1.1, this has proven to be a difficult task for most companies. This thesis will attempt to help companies to get a better understanding of mobile commerce and to increase the effectiveness of their m-commerce activities. Firstly, it is crucial to understand why consumers behave in a certain way and why they make or do not make certain decisions in the m-commerce environment. Therefore, the first part of the literature review will focus on thoroughly elaborating current issues in m-commerce, especially usability related issues. Possible solutions for these issues will also be discussed in this part.

Secondly, it is important that this knowledge about customers is applied to the design of the mobile interface. Based on literature review of user interface design and design aesthetics (2.3 & 2.4), combined with an analysis and comparison of Asian and western application design (3.1), certain variables are identified that can impact perceived usability and willingness to buy. Subsequently, these variables will be tested and analyzed in order to find the most optimal aesthetic design of mobile websites and applications. The combination of both the lessons learned from prior studies and the findings of the conducted experiment will provide valuable information for companies on how to optimize their mobile interface, taking into consideration the current issues in m-commerce. With this information, companies can increase traffic, loyalty, conversion rates and ultimately the profit of their mobile commerce activities.

Next to the relevance for companies, this topic could also provide valuable new additions to prior research in the field of m-commerce. Until now the subject of user interface design has been studied by a broad amount of researchers. Lohse and Spiller (1998a) provided a useful framework to effectively design a user interface for internet retail stores. They suggested that this framework could be used as checklist for online retail store designers. Kim and Eom (2002) built upon this framework and added a number of extra guidelines to the framework. Next to Lohse and Spiller's guidelines, the guidelines provided by Shneiderman (1998) has long been considered as standard for optimizing the user interface design. However, the above mentioned framework and guidelines are specifically designed for e-commerce. Considering the fact that m-commerce and e-commerce differ significantly in their dynamics, it is difficult to apply these recommendations to m-commerce.

In contrast to e-commerce, research has provided limited findings on user interface design for mobile devices. Gong and Tarasewich (2004) applied numerous of Shneidermans guidelines, complemented with mobile device specific guidelines, in order to develop a user interface design checklist for mobile retailers. Lee and Benbasat (2003) provided a framework for mobile interface design, based on seven design elements of m-commerce interface (7Cs). Although these studies provide useful information on the subject, the

findings of these studies can't be characterized as profound conclusions. The studies are rather a first attempt to create a solid framework that could be broadly used by mobile interface designers. To be able to develop such a framework, in-depth research has to be conducted on the various components of mobile interface design.

It has already been mentioned that one of those components, the design aesthetics of mobile websites and apps, has gained an increasing attention lately. Prior studies have focused on the effect of design aesthetics on various types of emotional reactions, most prominently the effect on perceived usability (Cyr et al., 2006; Hartmann, et al., 2007; 2008; Sauer & Sonderegger, 2009; Sonderegger & Sauer, 2010). The findings of these studies provide interesting findings on the effects of design aesthetics, which has proven to significantly affect a considerable amount of emotions.

However, in these studies design aesthetics are tested as a conjoint factor, not revealing the separate effect of specific variables of design aesthetics. Although it has to be noted that in various studies a distinction is made between classical and expressive aesthetics, there is no prior research in which the effect of separate specific variables of design aesthetics is conjointly examined. There are researchers who conducted studies on examining the effect of one of those specific variables though, for example the effect of symmetry (Bauerly & Liu, 2006; Tuch et al., 2010). However, in these studies only one separate variable is tested. In order to develop an aesthetically optimized user interface and to make recommendations based on this optimization, a conjoint analysis of the separate variables is needed to reveal the influence of the single variables. This is a gap in the research on m-commerce that still has to be filled and this thesis attempts to do so. Guidelines for the aesthetical design of a mobile interface could be the first step towards an extensive and consistent set of guidelines for mobile designers that covers every component of user interface design.

2. Literature review

This chapter will focus on outlining previous research in the field of mobile commerce. The chapter is organized as follows: Firstly, the concept of mobile commerce is defined. Secondly, current issues faced by companies in the m-commerce environment are discussed and possible solutions are proposed. In the third part, prior research of user interface design, existing guidelines for both e-commerce and m-commerce and prior research of design aesthetics are discussed. In the fourth and last part of this literature review, eight different variables and the underlying theory are discussed. Based on this theoretical foundation, hypotheses are developed and the underlying conceptual framework is presented.

2.1 Defining Mobile Commerce

The figures presented in the introduction show that mobile commerce has unarguably become an important part of companies' marketing activities and even holds the potential to surpass e-commerce in the upcoming decades. In order to fully understand mobile commerce, its dynamics and its potential, it is important to have a clear and consistent definition of what m-commerce exactly is and what activities are considered to be m-commerce activities. In existing literature however, there have been some issues regarding this exact definition. Some studies suggest that mobile commerce can be viewed as being a subset of e-commerce (Coursaris & Hassanein, 2002; Stafford & Gillenson, 2003). Based on this view, m-commerce can't be considered as a completely separate marketing channel. However, other studies point out the unique characteristics of mobile commerce, which make m-commerce such distinctive from e-commerce that it can be viewed as a separate channel (Balasubramanian et al., 2002; Shankar & Balasubramanian, 2009; Shankar et al., 2010). In this study, mobile commerce is assumedd to be a separate channel, next to the traditional brick-and-mortar channel and the e-commerce channel.

The most important characteristics that distinguish m-commerce from the other channels are characteristics related to mobility and accessibility (Shankar & Balasubramanian, 2009). With the emergence of m-commerce, retailers now can enter the consumer's environment through the mobile device practically anywhere, anytime. A characteristic that strengthens this is the personal nature of mobile devices (Shankar et al., 2010). Based on these unique characteristics and dynamics, a formal definition can be developed. In this study, mobile commerce is defined as the two-way or multi-way communication and promotion of an offer between a firm and its customers using a mobile medium, device, or technology (Balasubramanian et al., 2002; Shankar & Balasubramanian). An important addition to this definition is that this study focuses solely on smartphones as being the used mobile device. Other devices, such as tablets, smartwatches and e-readers, are disregarded in this study. Furthermore, it has to be noted that according to this definition, laptops and notebooks can also be considered to be mobile devices as they don't require a fixed network anymore these days. However, these devices are considerably larger than the abovementioned devices, which makes them unfit for use in specific situations and makes them less portable. This is

contrast with the unique characteristics of mobile commerce that were identified before. Therefore, laptops and notebooks are considered to be devices belonging to the e-commerce channel. In case the term 'desktop' is used in this study, which symbols the e-commerce channel, it can also be replaced by the words laptop or notebook as they are regarded as similar in this case.

2.2 Issues in M-Commerce

In order to increase the adoption rate and effectiveness of mobile commerce, it is essential to understand the underlying issues that deter consumers from engaging in m-commerce activities. Prior research identified a broad number of issues that companies have faced or still face nowadays. Ultimately, the goal of identifying these issues is to tackle them and increase the global adoption of mobile commerce. Davis' Technology Acceptance Model (TAM) identifies perceived ease-of-use and perceived usefulness as the two key factors that influence adoption of a technology (Davis, 1989). However, these two factors are at their turn influenced by a various amount of other factors. The aforementioned identified issues in m-commerce (section 1.1) are related to these factors, with usability being the most important issue. Therefore it is important to discuss the current issues and look into possible solutions, allowing m-commerce to be globally adopted.

As the identified issues include various aspects of m-commerce, the issues are classified into different categories. In prior research, several studies attempted to develop a consistent classification system for m-commerce issues (Tarasewich, Nickerson & Warkentin, 2002; Tarasewich, 2003a; Shankar & Balasubramanian, 2009; Shankar et al., 2010). However, these classification systems differ significantly from one other. This could be explained by the fact that many issues overlap each other. For example, compare weather issues and network issues. In the case of bad weather, one might not be able to use his device properly. On the one hand one could say that this is an external, environmental issue since the weather can't be influenced. On the other hand though, one could say that mobile technology isn't advanced enough yet to make a mobile device suitable for every weather condition. This would make it a technological issue.

For the sake of simplicity, the issues are classified into three different categories: usability, technology and privacy and security. Since usability related issues has been identified as the major issues in m-commerce and the fact that this is the main issue upon which this study is build, the other two issues will not be discussed in this chapter. However, an extensive review of these issues and possible solutions can be found in Appendix D. The three categories are defined as follows:

- *Usability:* This category revolves around the question whether a user is limited in using a retailer's mobile application or browser on a given time and location, performing a certain action. These issues could also be described as context issues (Tarasewich, 2003a). In this category, two assumptions are made. Firstly, it is assumed that the user experiences no technological failures, i.e. the mobile application or browser can be accessed and works exactly as intended. Secondly, it is assumed that the user is not hindered in performing any actions due to privacy, security or trust issues.
- *Technology:* This category revolves around the question whether a user is limited in using a retailer's mobile app or browser due to technological limitations or failure. As mentioned before, almost all issues in m-commerce can in one or another way be explained by limitations in technology. However, a distinction can be made between issues with the current technology and issues that could potentially be fixed by a more advanced, not yet existing technology. Throughout the different categories, the latter will be discussed because it provides possible solutions for current issues. However, in this category only the issues with the current technology will be discussed.
- *Privacy and Security:* Next to context issues and technological issues, privacy and security related issues can also hinder a user in using mobile applications or browsers or in performing certain actions (e.g. paying for a product on a mobile device). In this category these issues are discussed.

2.2.1. Usability

Mobile devices distinguish itself from desktops by three important characteristics: portability, location sensitivity and untetheredness (Shankar & Balasubramanian, 2009). Although these characteristics create many new business opportunities, it also creates many challenges to deal with. Whereas traditional desktop computers are stationary and only have to deal with a limited amount of contextual concerns, mobile devices face significantly more contextual challenges (Tarasewich, 2003a). Mobile devices nowadays have become a cultural object; it has created a mobile lifestyle in which consumers routinely use their mobile device for many different activities (Shankar et al, 2010). This implies that the consumers use their mobile devices in a various sets of time, location and activities that are being performed at the given time. The environments in which people use the application are relatively unstable from one moment to the next. As a result of this, many external factors could influence or hinder consumers in using their mobile device. In this section, these different issues are discussed. The issues can be divided into three categories: Environmental issues, user characteristics and physical limitations.

2.2.1.1 Environmental issues

When browsing online, a consumer is faced with many activities that require attention (e.g. social media and mail notifications, advertisement pop-ups, browsing among several tabs). In a traditional, wired e-commerce environment the environment outside the web is relatively stable. This implies that the consumer can devote a consistent amount of attention to performing tasks on the computer (Tarasewich, 2003a). However, when it comes to mobile e-commerce, this environment changes dramatically. The mobile environment is typically an unpredictable, unstable environment that can change within a matter of minutes. The outside environment often limits the users in paying full attention to their device.

One factor that can hinder a consumer from using his device is the weather condition (Tarasewich, 2003a; Duh, Tan & Chen, 2006). Although some tech companies have developed waterproof smartphones in the past years (Digital Trends, 2016), the majority of consumers doesn't own a waterproof phone. Moreover, rain can still pose a problem in terms of visibility. In the opposite case, with bright sunlight, visibility can also be limited. The majority of smartphones nowadays offer the option to adjust the screen brightness to enhance the visibility in these conditions, but that still may not always solve the problem.

Since the outside environment will always be unpredictable, the only possible solution to tackle this problem is to develop more advanced technological devices that are more suitable for any weather condition. One way to do this is to further develop the visibility qualities of the screen. Another option, which already exists for some years, is the use of a personal voice assistant. In 2011, Apple released the first personal voice assistant, called Siri. Over the years, the intelligence of the application has grown significantly and the abilities of the app are broad (O'Boyle, 2015). Further development of these personal voice assistants could eliminate the amount of physical (i.e. using hands) effort in using a mobile device. However, it seems unlikely that devices will be completely audio controlled in the future, since smartphones consist of two parts: audio and visual (Shankar et al, 2010). In particular when it comes to buying products online, the visual aspect is important.

Next to that, there is another important environmental factor that could hinder consumers in using their device. People nowadays use their smartphone at any place and time. Researchers even found a trend that people start taking the public transit more often so they can use their smartphone (Brown, 2015). The spaces in which consumers use their device are often crowded places, like streets, public transit, shopping malls, etcetera. These places generate a lot of possible distractions, which often hinder a consumer from using their phone. Noise could hinder people from using audio input and output, while the crowd around users requires constant attention from the user. If one doesn't pay attention, he could easily bump into another person because of being too focused on his device. Furthermore, consumers often use their phone while being engaged in social activities which forces them to multitask, leading to a decreased attention to their phone (Dunlop &

Brewster, 2002). Possible solutions for these environmental issues are to develop context-aware systems with minimal attention interfaces (Tarasewich, 2003a) and an interface that supports task interruptions (Dunlop & Brewster, 2002).

2.2.1.2 User characteristics

This year, global smartphone usage is expected to hit two billion (statista, 2014). With this many users, spread across the entire globe, consumers are completely different from another. In the early years of m-commerce, smartphones were mainly used by Asian and Western young adults. However, over the past years smartphone usage has increased rapidly. Correspondingly, the variety of the consumers has widened. An increasing amount of older adults (50+) are owning a smartphone these days. This can be explained by the fact that older adults experience more difficulties in maintaining their social life (due to physical or mental conditions). A smartphone facilitates easier access to their relatives and friends (Anderson, 2015a).

In addition to that, there is a rapid increase of smartphone ownership in developing countries. Ownership in these countries nearly doubled since 2013 (Poushter, 2016). This can be explained by two things. Firstly, many of these countries are experiencing a rapid economical growth, increasing the wealth in these countries. Therefore, more people are able to buy a smartphone (Boshoff, 2014). Secondly, device manufacturers developed and introduced affordable devices (i.e. priced below $100) in the emerging markets. These low-priced smartphones enables customers in these markets to own a phone, despite their relatively low income (Boshoff, 2014; Tshabalala, 2015).

These trends have led to a user base that differs significantly in terms of geographical, demographical, socio economic and behavioral characteristics. Companies need to take these differences into account when targeting their customers. However, with such a diversity, issues are inevitable. One major issue is the adoption of mobile technology by older adults. Although an increasing amount of older aged persons own a mobile device these days, the adoption rate is still far behind the rate of other age groups. Especially adults aged 65 or older show a very low adoption rate (Anderson, 2015b). This is caused by several factors.

One factor is that older people are facing physical challenges when using new technology. Many older customers suffer of physical health condition that makes it difficult or challenging to read (Smith, 2014). The limited screen size limits their options to participate in the mobile medium (Shankar et al, 2010; Tarasewich, 2003a). Unfortunately, it seems to be difficult to solve this issue. Screens will probably not significantly increase in the nearby future. Therefore, the best solution for companies in this case is to target these type of customers via desktop or tablet (which technically also belong to mobile devices). A study of Smith (2014) supports this solution, as he found that seniors are more likely to own a tablet or e-book reader than a smartphone.

A second factor that hinders senior adults from adopting mobile technology is that they simply refuse to use it because they are skeptical about the benefits of technology (Bigné, Ruiz & Sanz, 2007; Smith, 2014). However, Smith (2014) also found that once seniors adopted digital technology, it often becomes an integral part of their daily lives. Next to that, a third factor that hinders seniors from adopting technology is the lack of skills and knowledge to use new technologies. A significant majority of older adults state that they need help to get to learn new technologies (Smith, 2014). This is supported by findings of Bigné, Ruiz and Sanz (2007), who suggest that it is necessary to take into account the user's capacity to understand the changes and complexity that arise from adopting a new technology. In contrary to the first issue related to age, the issues mentioned above can be solved. First of all, it is important that senior adults are properly informed about the advantages of adopting mobile technology, which will take away their skeptical attitude towards new technologies. Moreover, senior adults should receive some form of training to understand the basic principles of new technologies.

Lack of skills and knowledge to use new technologies is not merely an issue among seniors. The rapid development of mobile technology has made the smartphone a device with many functionalities. However, many people are not aware of all the possibilities their smartphone gives them or they simply lack the skills to fully operate their device. Mobile device manufacturers should bear this in mind and design the devices for a widespread population (Dunlop & Brewster, 2002). The focus in designing devices should lie on simplicity of the device. The more simple a device is, the more people are able to use it. The same applies to retail companies when they design their applications or mobile browsers. Furthermore, mobile device manufacturers and companies should pay attention to educating their consumers about all the possibilities of mobile technology (Shankar et al, 2010).

Other characteristics that could influence m-commerce are characteristics related to socioeconomic status and culture. Smartphones have long been a status symbol. They were expensive items, which means they were only accessible for the more wealthy people among us. Nowadays however, smartphones are part of our daily life. This doesn't mean they are cheap though. According to a study of Anderson (2015b), there are significant differences in smartphone ownership between low and high household incomes. Not surprisingly, this study also found that highly educated people are more likely to own a smartphone than low educated people. This implies that, at least for some part, this could be explained by the fact that these people can't afford a smartphone. The same applies to emerging markets, in which many people still don't have the resources to buy a mobile device. The solution for this issue has already been mentioned in the introduction of this section: introducing low-priced smartphones to make smartphone ownership more accessible for a larger population. Next to that, if the emerging markets continue to grow at the current pace, the smartphone market will grow simultaneously in the upcoming years. Lastly, differences in culture should be taken into account when targeting specific

customers, since studies found that culture could influence consumer's intention to use mobile commerce (Dai & Palvia, 2009).

It is evident that no customer is equal, all the user characteristics form unique customers. The problem is that for optimal m-commerce results, customers need to be targeted based on their customer profile. With personalized offers and settings, m-commerce can be very effective. The challenge though is to combine all these different characteristics into one customer profile. The solution for this problem is to be found in the emerging field of data driven marketing. An increasing amount of companies collect all sort of data of their customers in order to create an accurate customer profile. All this data is analyzed and used to make m-commerce as personalized and effective as possible. However, data driven marketing brings up questions and discussion about what is legally and ethically acceptable. These issues will be discussed in the privacy and security section of Appendix D.

2.2.1.3 Physical limitations

In the previous section, it has already been discussed that senior adults cannot fully participate in mobile technology due to physical and health issues. This is a specific issue for that age group and therefore it is not discussed in this section. However, senior adults are not the only age group that experience physical limitations. Customers are often physically limited to use a device due to a certain action that is performed at that time, like walking or driving a car. This section will further elaborate on these physical limitation issues.

This entire issue revolves around one underlying trade-off: The trade-off between usability and portability (Tarasewich, Nickerson & Warkentin, 2002). Larger screens and higher screen resolution enhance the usability of a mobile device. However, this reduces the portability of the device. Firstly, it is more difficult to carry a smartphone with you when the size is large. Secondly, larger screen size and higher screen resolution decreases the battery life of a device. This means it has to be charged more often and therefore becomes less portable.

Companies constantly need to take this trade-off into account when developing and designing new smartphones. For many years there was a clear shift towards portability, with devices that continued to shrink in both size and weight (Tarasewich, 2003a). However, in the last decade the size of the average smartphone has nearly doubled (Barredo, 2014). Nowadays, the average size has even increased that far that researchers came up with a new name for certain devices: phablets. These devices can be considered as too big to be called a phone, but too small to be seen as a tablet. However, it is expected that the average size of smartphones will start to decrease again in the upcoming years (Bonnington, 2013).

The constant change of size of smartphones creates several issues. Smartphones are designed for one-handed interaction. It is designed in such a way that the consumer can perform any interaction using just one hand, leaving the other hand free to carry items or, if needed, to assist the other hand when typing a message (Samet et al, 2011). However, with the current size of smartphones, it is practically impossible to fully control the device by

using just one hand. This problem was already indentified in the time that device sizes were considerably smaller (Kim et al, 2002; Tarasewich, 2003a). Next to that, the device itself has become difficult to carry, since it is quite a challenge to put a device with almost the size of a tablet in your pocket. It's obvious that the large size can cause some serious trouble in terms of usability and portability. On top of that, the (touch)screen could cause even more phyiscal limitations. The typical mobile device is designed in such a way that it requires only one's fingers as natural input tool (Shankar et al, 2010). Touchscreens are highly sensitive and fingers often lack accuracy and precision, especially the thumb, which leads to large error rates (Samet et al, 2011).

To sum up the abovementioned issues: the current (average) smartphone is often too big to control by using just one hand, is difficult to carry and experiences large error rates due to touchscreen sensitivity and lack of finger accuracy. These issues already occur when a consumer is just standing or sitting (i.e. not moving). However, one key characteristic of smartphones is not taken into account in this case: mobility. Mobile devices are designed to be used while being on the move. As already discussed in the environmental issues section, this causes a lot of potential distractions and limitations, which increases the issues mentioned above. This is confirmed by a study of Kim et al (2002), which found that users experienced more difficulty using a website while moving compared to performing the same actions while not moving.

The question is whether these issues can be solved. For some part of the issues, technological improvements could be the solution. Touchscreens could be improved in terms of sensitivity and accuracy to reduce the large error rates. Furthermore, attention should be paid to minimizing needless features to further reduce error rates (Samet et al, 2011). The other usability and portability related issues in this section highly depend on the average size of a mobile device at that moment. Currently, smartphones could be reduced in size to make them easier to control with one hand and make them easier to carry. However, doing so implies that other issues arise. Smaller screens can be more difficult to read, which can lead to user frustration (Tarasewich, 2003a). There will always be a trade-off between usability and portability. According to the article of Bonnington (2013), screen sizes will decrease and eventually stabilize at mid-sized (3.5 to 4.9 inch) screens. In this case, the market mechanism is a big part of the solution; companies simply need to adjust the screen sizes according to this mechanism. In addition to that it is important that companies pay attention to the mobile interface design to improve the usability of mobile applications and browsers, no matter the size of the screen. If companies want to increase their mobile revenue, their app or mobile browser should in any case be usable for the customers.

Lastly, the fact that consumers use their device in all kind of situations causes another problem: safety risks. Although smartphones are designed to be used while being on the move, it doesn't mean that one can perform any action simultaneously while using their device. The best example of this is using a smartphone while operating a vehicle. Research

from AT&T (2015) shows that 7-in-10 people engage in smartphone activities while driving, with nearly 4 out of 10 of the respondents reporting to tap into social media while driving. Not surprisingly, it is found that engaging in smartphone activities significantly affects user's attention while driving and can therefore be considered as a big threat to road safety (Nunes & Recarte, 2002). According to the CDC (Centers for Disease Control and Prevention), in the U.S. alone each day 8 people are killed and over a thousand people are injured in crashes that involve a distracted driver. It is evident that this is a big issue. One way to solve this safety issue is to implement strict legislation about engaging in smartphone activities while driving to enforce the safe use of m-commerce devices (Tarasewich, 2003a). Another solution is to develop technological improvements so that users can use their device without being distracted (e.g. handsfree carkits). The last solution is to clearly inform users about the dangers of using their phone while driving.

2.3 User Interface Design

In the previous paragraph, it has become clear that there are still many issues to be solved in the field of mobile commerce. However, many of these issues (technological, security, privacy, environmental) can be solved by improvements in mobile technology, like the 5G network. Other issues though require considerably more attention in order to be solved, especially usability related issues. Proper user interface design has the ability to significantly improve the usability of handheld devices. Although some studies proposed guidelines for mobile devices in the early years of m-commerce, these guidelines are not complete guidelines that can be used as unified standard by mobile interface designers. Moreover, these guidelines date from over a decade ago and are therefore outdated. In the first section of this paragraph, several of the abovementioned guidelines from both e-commerce and m-commerce are discussed. This is followed by a review on design aesthetics, an important part of the interface design and the core of this paper. In this second section, the concept of design aesthetics is defined and prior research in this field is discussed, along with two measurement models.

2.3.1 Existing User Interface Guidelines - E-commerce

Ever since the emergence of e-commerce, user interface design has been a topic of interest. Studies have found that an effective design of user interfaces (Baty & Lee, 1995; Lohse & Spiller, 1998a; Lohse & Spiller, 1998b) and online store atmosphere (Eroglu et al., 2003) can significantly increase an online store's traffic and sales. Based on this knowledge, several scientists developed user interface design guidelines for e-commerce. The two most well known guidelines, those of Lohse and Spiller (1998a) and Shneiderman (1998) are briefly discussed below.

2.3.1.1 Lohse & Spiller's guidelines

In their study, Lohse & Spiller (1998a) identified 32 online retail store attributes, which they categorized into five groups: merchandise, service, promotion, convenience and (other) interface variables. Findings of their study showed that improving browsing and navigation capabilities and especially product lists information could significantly improve online retail stores' traffic and revenue. On the other hand, they found no significant results for variables that created a 'fancier' user interface, such as image size and background patterns. Lohse and Spiller suggested that their list of features and attributes could be used as a checklist for online store designers, in combination with the findings of their study. Kim and Eom (2002) added a couple of variables to these proposed attributes (currency of information, privacy and security) to expand Lohse & Spiller's research. They found that convenient and dependable shopping, reliability of the retailer, additional information and product perceptions were significant factors to the satisfaction of online shoppers.

2.3.1.2 Shneiderman's guidelines

In 1998, Ben Shneiderman wrote a book on how to design the user interface. He attempted to create unified guidelines that could be used by any online retail store designer in the world. In his book, Shneiderman developed the so-called "Eight Golden Rules of Interface Design", based on knowledge of the past decades. These guidelines are still applicable to e-commerce interface designs nowadays. The eight rules developed by Shneiderman are:

1. Strive for consistency
2. Cater to universal usability/Enable frequent users to use shortcuts
3. Offer informative feedback
4. Design dialogs to yield closure
5. Prevent errors
6. Permit easy reversal of actions
7. Support internal locus of control
8. Reduce short-term memory load

Although not every part of user interface design is included in these guidelines, following these eight golden rules should provide designers with the tools to build an interface of sufficient quality. However, the guidelines proposed by Shneiderman and Lohse and Spiller were created in a time that mobile technology was in a very early stage. Considering the very different characteristics and dynamics of m-commerce, these guidelines are difficult to apply to mobile interface design.

2.3.2 Existing User Interface Guidelines - M-commerce

Because of the fact that the abovementioned guidelines are not perfectly suited to be applied to mobile commerce, new guidelines focused on mobile devices had to be developed. Over the past decade, several studies attempted to develop such guidelines,

covering as many aspects as possible. The most complete set of guidelines currently in existence is the one proposed by Gong and Tarasewich (2004). They based their guidelines on the eight golden rules of Shneiderman (1998). According to Gong and Tarasewich, four of the eight guidelines (Offer informative feedback, support internal locus of control, design dialogs to yield closure and enable frequent users to use shortcuts) could be carried over to mobile devices without explicit changes. The other four guidelines required modification in order to make them applicable to mobile devices. Next to these eight guidelines, the study proposed seven additional guidelines specifically for mobile device interface design. Combined, these 15 guidelines could serve as a solid checklist for mobile interface designers.

In another study, Chan et al. (2002) developed eight guidelines for mobile interface design based on the findings of their own research. In their research, various predefined tasks were performed on specific predefined websites, like booking a flight on the website of an airline company. The eight guidelines that followed were mainly guidelines to improve the usability of mobile websites and applications. Most of the guidelines were related to the navigational structure of a website, which obviously is an important factor in the user interface design, especially in the case of mobile commerce with its typical small screen sizes. Peter Tarasewich (2003a; 2003b) also focused his studies on enhancing the usability of handheld devices. His findings outlined the importance of context in mobile interface design. Hence, the recommendations in his studies mainly focused on how to deal with contextual concerns, like influence from the user's environment.

Venkatesh, Ramesh and Massey (2003) compared the importance of usability attributes between e-commerce and m-commerce. Their findings showed significant differences between Web-based and wireless sites. Especially the factors ease-of-use and made-for-the-medium proved to be significantly more important factors in the wireless context, whereas content was important in both Web-based and wireless sites. Based on this results, the researchers provided several recommendations to effectively design mobile interfaces. Their most important recommendations included possible improvements for and simplification of the navigation structure of mobile browsers. Furthermore, their results strongly suggested that the ability to present content to users in a customized fashion is a key to success in the wireless context.

Other studies developed frameworks for the effective design of mobile interfaces, from which certain guidelines could be derived. One of those frameworks was proposed by Lee and Benbasat (2003; 2004). They built their framework upon Rayport and Jaworski's (2001) seven design elements of an e-commerce interface design, the so-called 7C's (context, content, community, customization, communication, connection and commerce). In addition to these 7C's, they identified two M's in order to specify the framework on mobile commerce: mobile setting and mobile device constraints. The combination of the 7C's and 2 M's were proposed as a new framework for m-commerce. Based on this framework, the

authors suggested a list of interface implementations/guidelines to improve the mobile interface design.

Another framework was proposed by Adipat and Zhang (2005). The core of their framework consists of four major components: user, context, information presentation, and data entry methods. These components are further divided into major issues. Based on these components and issues, the authors proposed several guidelines and recommendations in order to help mobile designers with designing an effective mobile interface.

It can be concluded that many studies have attempted to develop a set of guidelines or a framework for mobile interface design. Although the findings of these studies often recognized the same issues and had some degree of similarity in their recommendations, none of these guidelines or frameworks can be considered as a unified standard in mobile commerce. In the past decade, studies in the field of mobile interface design shifted its focus towards interface design for specific groups that required an adjusted interface, like blind people (Kane, Bigham & Wobbrock, 2008; Krajnc et al., 2011), elderly people (Kobayashi et al., 2011) and low-literate people (Chaudry et al., 2012). These studies attempted to fill some existing gaps in the current mobile interface literature. Typically, the existing mobile interface guidelines and frameworks are focused on usability, reliability and personalization. However, the influence of graphical and emotional content and design on mobile browsers and applications is often neglected in these studies. In the last years however, an increased scientific interest in this so-called design aesthetics can be identified.

2.3.3 Design Aesthetics

Design aesthetics (DA) in this study can be defined as the balance, emotional appeal or aesthetic of a mobile website or application, which may be expressed through colors, shapes, font type, music or animation (Cyr et al., 2006). In the nineties, studies on user interface design mainly focused on increasing the usability of the interface design. In the aforementioned study of Lohse and Spiller (1998a), the authors included store presentation variables (image size, background patterns, number of buttons on storefront) as interface design features that have a possible impact on store traffic and sales. However, they found that these variables did not affect traffic and sales. They suggested that interface designer should rather focus on navigation features, as consumers want to find products quickly and effortlessly. Their conclusion on this matter was that no amount of 'sparkle' in the presentation of products can overcome a site design with poor navigation features. On the contrary, Kurosu and Kashimura (1995) found evidence that the aesthetic aspect of the interface could strongly affect the user when they try to evaluate the interface in its functional aspects, suggesting that designers should not merely focus on improving the inherent usability of the interface.

In the beginning of the first decade of the 21th century, several studies confirmed Kurosu and Kashimura's findings that other design features than (inherent) usability play an important role in determining user satisfaction (Norman, 2004; Tractinsky, Katz & Ikar, 2000). Over the past decade, these findings has led to a continuously increased interest in design aesthetics. In these years, studies have attempted to find the effect of design aesthetics on various factors. One effect is the Aesthetic-Usability Effect. This effect describes the phenomenon in which users tend to perceive more aesthetically appealing interface designs as easier to use than less appealing designs. This effect has been observed in several experiments (Cyr et al., 2006; Hartmann, et al., 2007; 2008; Li & Yeh, 2010; Sauer & Sonderegger, 2009; Sonderegger & Sauer, 2010), which all found that design aesthetics significantly impacted perceived usability (PU) and perceived ease-of-use (PEOU). In addition to this, some of these studies found that PU and PEOU had a significant impact on customers' loyalty (Cyr et al., 2006) and trust (Li & Yeh, 2010). The latter study even found a direct significant influence of design aesthetics on customers' trust in m-commerce.

Although the Aesthetic-Usability Effect is widely observed and supported, some studies have identified limitations of the effect. Hartmann, Sutcliffe and de Angeli have conducted two experiments (2007; 2008) in which they found evidence of the existence of the A-U effect, but only partially supported the effect. They discovered a framing effect in which the user's judgement of aesthetics and overall preference varied according the question. Without the framing question, users rated the aesthetically more appealing website to be superior on aesthetics and usability, even though the objective evidence of usability problems did not support this perception. The authors concluded that this was caused by a halo effect, which implies that high perceived aesthetic appeal of a website can override users' poor usability experience. These findings are completely in contrast with the findings of Lohse and Spiller (1998a), who claimed that an aesthetically appealing design can't overcome a site design with poor navigation features.

However, when users were given a question framed for serious use, their preferences shifted towards the design with high objective usability, showing that the correlation between design aesthetics and usability isn't as simple and straightforward as has been suggested in other studies. Van Schaik and Ling (2009) has found similar framing effects, suggesting that context (mode of use) is a crucial factor influencing the stability of users' perception: users tend to change their preference if context (e.g. performing a specific task on a website) is provided.

Next to the Aesthetic-Usability effect, studies found that design aesthetics significantly influence users' emotions. Sauer and Sonderegger (2009) found an effect on emotions in general, whereas other studies found significant effects on specific emotions or specific groups, like the emotional reaction of males (Nanda et al., 2008), perceived online service quality (Wang et al., 2010) and trustworthiness of a website (Kim & Moon, 1998). Cyr (2008) found that website design has a significant impact on users' perceived trust and satisfaction,

which ultimately lead to a higher loyalty of the user. However, the findings of this study also suggest that this effect could be influenced by a user's culture, as in some cultures not all these effects were found to be significant. Another effect observed in the aesthetics literature is the amelioration effect of visual design and aesthetics on content credibility (Robins & Holmes, 2008). This effect describes the phenomenon in which users prefer content with a higher aesthetic treatment over content with low aesthetics treatment, while the content in both cases was exactly the same.

Measurement models

Despite the numerous studies in the field of design aesthetics, only a limited amount of studies have attempted to develop a measurement model of users' perceptions of website aesthetics. To date, only two studies have succeeded to develop a consistent model with high validity. The first model was developed by Lavie and Tractinsky (2004) and has often been used as base for other studies in the field of design aesthetics. In their study, the authors found that users' perception of aesthetics consists of two dimensions. The first dimension is termed "classical aesthetics" and the second dimension is termed "expressive aesthetics". Classical aesthetics relate to the aesthetic notions that emanate from ancient times. These notions highly value a design that is simple, orderly and clear and can be compared to variables related to objective usability. On the contrary, expressive aesthetics can be defined as aesthetics that stem from the designer's creativity, originality and expressive power. This dimension of aesthetics typically enhances the users' subjective usability by evoking emotions.

The second model, develop by Moshagen and Thielsch (2010), used the model of Lavie and Tractinsky as base and attempted to find more than two dimensions in aesthetics. In their model, the Visual Aesthetics of Website Inventory (VisAWI) model, the authors identified four interrelated facets of perceived visual aesthetics: Simplicity, Diversity, Colourfullness and Craftsmanship.

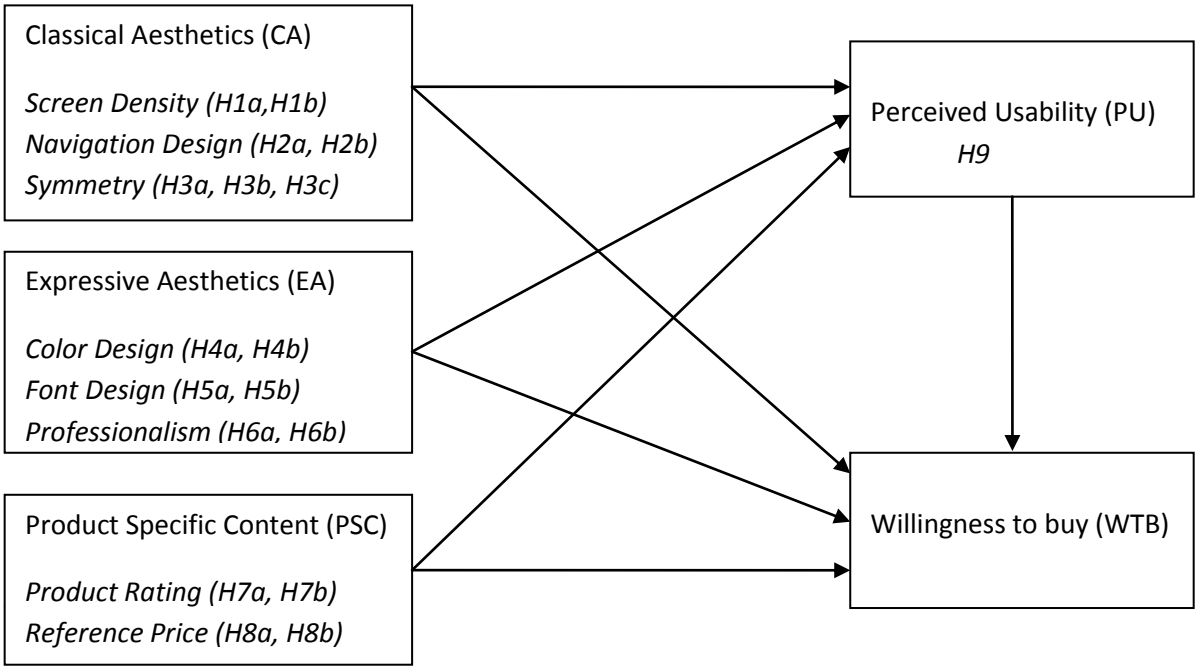
- *Simplicity* refers to aspects related to the figural goodness concept, such as balance, structure and orderliness.
- *Diversity* includes aspects related to variety, visual richness, dynamics and creativity.
- *Colourfullness* reflects the selection, placement and combination of colors.
- *Craftmanship* relates to whether the site was designed with skill and care using modern technologies.

The combined findings of abovementioned models will be used as base for the experiment in this study, in which both the effect of classical and expressive aesthetics will be tested. The variables used in this study are highly related to variables used in the classical-expressive model and the VisAWI model.

2.4 Conceptual framework and hypotheses

In this thesis, it will be examined what the effect of design aesthetics is on both perceived usability and willingness to buy. Next to that, the effect of product specific content on perceived usability and willingness to buy will be examined. A graphical depiction of this framework is shown below, representing the relationship between classical aesthetics, expressive aesthetics, product specific content, perceived usability and willingness to buy.

Figure 2: Conceptual framework



2.4.1 Classical Aesthetics (CA)

Classical Aesthetics include aesthetic aspects that are concerned with the simplicity and functionality of the mobile browser or application (Lavie & Tractinsky, 2004). This is consistent with the simplicity facet in the VisAWI model of Moshagen and Thielsh (2010). The aspects of CA may also be considered measures of objective usability since they examine actual user behavior (Sauer & Sonderegger, 2009). For many years the CA aspects of interface design have been regarded as being the most important factors of user interface design (Childers et al., 2002; Lohse & Spiller, 1998a; 1998b; Park & Kim, 2003). However, as discussed in the previous paragraphs, in the last decade expressive aesthetics has also been identified as a key factor in designing user interfaces. Nevertheless, more recent studies do still regard classical aesthetics as a vital part of user interface design (Cyr, 2008; Hartmann et al., 2007; 2008; Koo & Ju, 2010; van Schaik & Ling, 2009). As mobile screens are considerably smaller than desktop screens, a simple and effective interface design could be even more

important for smartphones. This is confirmed by a study of Choi and Lee (2012), who found that simplicity is a major factor in mobile interface design. Additionally, they suggested that a clean, modern and balanced arrangement of graphical and textual items should be the main design motto of mobile interface design.

One CA aspect that will be examined in this thesis is the influence of screen density on both perceived usefulness and willingness to buy. In this case, screen density is defined as the amount of content shown on the screen. Little content implies a low density, much content implies a high screen density. Since the screens of online retailers are mainly made up of product information, the amount of products shown on the screen will either increase or decrease screen density. Consequently, a low screen density (less products featured on screen) implies that the product images shown on screen are larger than in the case of high screen density. Detenber and Reeves (1996) found that image size positively affects the arousal and dominance dimensions of emotional responses. According to their study, large images elicit stronger feelings than the same picture in a smaller size. In an article on his website, conversion expert Jeremy Smith (2014) suggests that larger images could lead to significant higher conversion rates, although this was tested for e-commerce and not confirmed for m-commerce.

Furthermore, studies haven't found any evidence that showing additional products enhances site traffic and sales (Lohse & Spiller, 1998a). On the contrary, Bauerly and Liu (2006) found that a large number of visual groups (in this case pictures) has negative effects on a website's aesthetics appeal as the website is perceived complex and overcrowded. This ultimately leads to lower perceived usability. Although their study is based on analysis of websites, it is assumed that the same applies for mobile browsers and applications, which even stronger rely on simplicity. Furthermore, it is expected that low screen density and therefore larger image sizes lead to a higher willingness to buy.

H1a: Interface designs with low screen density will lead to higher perceived usability
H1b: Interface designs with low screen density will lead to higher willingness to buy

Another classical aesthetics aspects that will be examined is the effect of navigation design on perceived usability and willingness to buy. In previous studies it has been found that a clear and simple navigation design significantly affects online store traffic and sales (Lohse & Spiller, 1998a), perceived user satisfaction (Choi & Lee, 2012; Cyr, 2008), perceived trust (Cyr, 2008), perceived enjoyment (Childers et al., 2001) and perceived usability (Bachiochi et al., 1997; Chan et al., 2002; Childers et al., 2001; Fang & Holsapple, 2007). However, studies by Hartmann, Sutcliffe and de Angeli (2007; 2008) found that sites with higher objective usability aren't necessarily rated as superior in perceived usability. As this study only investigates the visual aspects of mobile interface design, it can't be examined whether the objective usability is increased by the navigation design. Therefore, in this case it will merely be tested whether navigation design influences perceived (i.e. subjective) usability. This will be examined by varying the number of navigation buttons shown in the screen (e.g. having

more extensive search options, having filter options, etc.). Hartmann, Sutcliffe and de Angeli (2007; 2008) also found that expressive aesthetics have a bigger impact on perceived usability than classical aesthetics. Next to that, Koo and Ju (2010) found that only the linking structure design of websites has significant impact on users' emotions and intention to buy, whereas no significant relations were found for menu design. In this study, a mix of menu buttons and linking buttons will be examined. Although it is not expected that a small difference in navigation design has a significant impact, it is assumed that considerable differences in navigation design of the mobile interface leads to higher perceived usability and willingness to buy.

H2a: Clear and simple navigation design positively affects perceived usability
H2b: Clear and simple navigation design positively affects willingness to buy

The last CA aspect that will be examined is the effect of symmetry on perceived usability and willingness to buy. The symmetry aspect has been tested in both the model of Lavie and Tractinsky (2004) and the model of Moshagen and Thielsch (2010). Although both studies found evidence that symmetry has an impact on users' perception of visual aspects, the importance of this aspect has found to be less important than other classical aesthetic aspects such as simplicity and orderliness. On the other hand other studies suggested that users consider symmetrical design as more appealing than an asymmetrical design, leading the users to have a significant preference for symmetrical designs over asymmetrical designs (Bauerly & Liu, 2006; Tuch et al., 2010). In the latter though, such a relation has only been found for male participants, whereas symmetry doesn't influence women's aesthetics judgements. Regarding the effect of symmetry on willingness to buy, no evidence was found that suggests a relationship between these variables. However, as it is believed that perceived usability significantly influences purchase intentions (Hall & Hanna, 2004; Lin & Wang, 2006), it is in this case expected that symmetry positively affects users' willingness to buy.

H3a: Symmetry positively affects users' perceived usability
H3b: Symmetry positively affects users' willingness to buy
H3c: Aesthetic judgements of male participants regarding symmetry are significantly different than those of female participants

2.4.2 Expressive Aesthetics (EA)

Expressive Aesthetics include aspects that are related to the expressive power, creativity and originality of the designer (Lavie & Tractinsky, 2004). In the VisAWI model, these aspects are captured in three facets of the model: Diversity, Colourfulness and Craftsmanship (Moshagen & Thielsch, 2010). Expressive aesthetics aspects are often associated with subjective evaluations of a website's usability and the emotions that are evoked by expressive aesthetics, which are typically stronger than emotions evoked by classical aesthetics (Hartmann, et al., 2007; 2008; Sauer & Sonderegger, 2009; Sonderegger & Sauer,

2010). As mentioned before, expressive aesthetics have long been neglected as major factor in designing user interfaces but is nowadays regarded as a key factor in user interface design. However, as mobile screens are considerably smaller and ask for a high level of simplicity to prevent overcrowding of the screen, the question is whether expressive aesthetic aspects are equally important in mobile interface design.

One expressive aesthetic aspect that will be examined is the influence of color design on perceived usability and willingness to buy. In the models of Lavie and Tractinsky (2004) and Moshagen and Thielsch (2010), colourfulness has proven to be one of the most important factors of perceived visual aesthetics of websites. It has also been found that preferred colors lead to higher ratings of aesthetic quality, which ultimately led to an increased purchase intention (Hall & Hanna, 2004). The influence of color has always been a topic of interest in science. Wilson (1966) found evidence of a U-shaped relationship between color wavelength and arousal effect, in which colors with more extreme wavelength are the most activating. In the U-shaped relationship violet/blue is at the extreme short wavelength end, with red/orange being at the opposite end of the U (long wavelength).

Findings of which extreme wavelength is the best to use in the retail environment are mixed. A study of Crowley (1993) suggested that activated consumers are more likely to engage in impulse buying and therefore suggested that in situations in which impulse buying is the goal, more extreme and activating colors such as red and blue should be used. Additionally, it was suggested that for retail environments where impulse buying of stylish merchandise is the goal, a red colored environment should be used since red is perceived as 'up-to-date'. In general though, Crowley (1993) found evidence that cool-colored (i.e. short wavelength, blue) store environments are preferred over warm-colored (i.e. long wavelength, red) environments. This is consistent with findings of other studies, who also found evidence that blue store environments positively affect purchase intentions (Babin et al., 2003; Bellizi & Hite, 1992; Gorn et al., 2004).

On the contrary, Wu, Cheng and Yen (2008) found that red color (rather than blue) has positive impacts on pleasure and arousal, which led to an increased purchase intention. An explanation for these conflicting findings can be a difference in culture. Kondratova and Goldfarb (2007) have found evidence that preferred colors differ across different cultures. These findings are confirmed by Cyr (2008), who found that users from more collectivist cultures such as China have a strong preference for visuals, whereas users from rather individualistic countries tend to prefer a logical and structured page lay-out. Next to that, colors have different meanings across cultures, with red being perceived as happiness in China, but danger in the United States. This could explain the contradictory findings of Wu, Cheng and Yen (2008), as their experiment was conducted in Taiwan. As the respondents in this study will mainly be western orientated, it is assumed that a blue color environment is preferred over a red environment. However, in this thesis blue colors will be tested against

neutral (i.e. black, white and grey) colors to test if the use of colors in general affects PU and WtB.

H4a: The use of extreme short wavelength colors positively influences perceived usability
H4b: The use of extreme short wavelength colors positively influences willingness to buy

The second expressive aesthetic aspect that will be tested in this study is the influence of font design on perceived usability and willingness to buy. In this study font design is defined as the variety of the used font in terms of color, size and style. Cyr (2008) categorized font design as a part of visual design aesthetics of a website. The results of this study implicated that visual design aesthetics have a significant influence on users' perceived trust, satisfaction and loyalty. However, the effect of font design on these emotions has not been tested separately. However, increased variety in font design may enhance the users' perception of visual aesthetics, according to the Diversity facet in the VisAWI model (Moshagen & Thielsch, 2010). Hill and Scharff (1997) found that the font style significantly impacted the readability of a website, which may lead to an increased usability of the website. They also found evidence that varying the combination of background/font colors could lead to different results in perceived readability.

In an attempt to extent this study, Hall and Hanna (2004) suggested that black/white font/background color combination should be used for educational sites, whereas for commercial websites more colored font/background combination should be used. This is consistent with the literature review on color design, which states that the use of colors leads to enhanced perceived visual aesthetic ratings and perceived usability. Furthermore, a study of Schmidt, Liu and Sridharan (2009) suggested that font size positively affects both ease of interaction and users' aesthetic preference, implying that bigger font sizes are better.

H5a: Variety in font design positively affects perceived usability
H5b: Variety in font design positively affects willingness to buy

The last EA aspect that will be examined is the effect of professionalism of the mobile browser or application on perceived usability and willingness to buy. In this case, there will be two extreme ends in terms of professionalism. On the one end, the site looks professionally, neatly designed with a relatively high use of lines to separate products and navigation buttons and the use of rather formal icons. On the other end, the design is creative and playful with a relatively low use of lines and rectangular shapes and the use of more creative icons. According to the models of Lavie and Tractinsky (2004) and Moshagen and Thielsch (2010), a more creative and original design is perceived by users as more visually aesthetic. Although the effect of a professional/playful appearance has barely been investigated, it is assumed that a more playful and creative look leads to increased perceived usability, based on the findings of the abovementioned models. Concerning the effect on willingness to buy, it is expected that more professional looking sites are considered as more

reliable. This implies that when a user is asked to buy a product, the user will prefer to buy the product on the more professional looking website. This would be consistent with findings of previous studies, in which preferences shifted as soon as users were given a task or additional context (Hartmann et al., 2007; 2008; van Schaik & Ling, 2009).

H6a: Professional appearance negatively affects perceived usefulness
H6b: Professional appearance positively affects willingness to buy

2.4.3 Product Specific Content

In addition to examining the effects of design aesthetics, the effect of showing product specific content on perceived usability and willingness to buy will be tested in this study. Product specific content in this context can be defined as product information that increases user's ability to compare products. This is achieved by testing two variables that are characterized by their referring/comparing abilities: external reference prices and product ratings. In the first case, users can take the external reference price as reference upon which they build their purchasing decision. In the second case, users can take given product ratings from an anonymous reference group as reference upon which they build their decision.

Although one could say that product ratings (i.e. showing yellow stars in this case) could have an effect on users' perception of visual aesthetics, it is hard to categorize PSC aspects as being aspects of design aesthetics. In the models of Lavie and Tractinsky (2004) and Moshagen and Thielsch (2010), no comparable aspects has been tested or mentioned as factors that could influence perceived aesthetic judgements. However, showing an external reference price and/or product rating could be considered as an increased informational value of a site's content and an increased variety of the site. According to the models, this could enhance perceived ratings of aesthetic beauty. Since more aesthetically appealing designs are preferred over less appealing ones, it is expected that both reference prices and product ratings increase the perceived usability of a mobile browser or application.

H7a: Showing product ratings increase users' perceived usability
H8a: Showing an external reference price increases users' perceived usability

Next to the effect of both variables on perceived usability, it will also be examined whether there is a relation between the two aspects and users' willingness to buy. In the case of the first aspect, product ratings, the hypothesis is based on the social influence theory. This theory was first proposed by Deutsch and Gerard (1955), who described two psychological needs that make individuals accept influence of others: normative (our need to be liked) and informational (our need to be right) social influence. In the latter case, individuals accepts influence of another individual or group as they believe the information obtained from them is evidence about reality. In addition to this theory, Kelman (1958) distinguished three processes of influence: compliance, identification and internalization. In case of compliance, an individual agrees to others because he or she hopes to achieve a favourable reaction from the group. Identification occurs when an individual changes his or her behaviour due to the

influence of someone that is liked. Lastly, internalization occurs when an individual accepts certain norms and values that are viewed as normal behaviour by groups or persons which are influential to the individual.

Regarding the effect of product ratings, a combination of informational influence and internalization could lead the user to adopt the opinions of the reference group about a certain product as being evidence about reality. This has been confirmed by a study of Burnkrant and Cousineau (1975), who suggested that people use others' product evaluations as source of information about a product. They found evidence that people perceive the product more favourably after observing that others evaluate the product favourably. In the absence of this observation, people perceive the same product as less favourable. In studies focused on social influence effects in an online environment, similar effects have been found with product ratings functioning as online sales assistants (Anderson, 1998; Chen & Xie, 2008; Sridhar & Srinivasan, 2012). Furthermore, evidence has been found that product ratings have the potential to significantly affect product sales (Chevalier and Mayzlin 2006; Dellarocas et al., 2007), although Moe and Trusov (2011) suggests that these effects are relatively short lived. Based on these findings, it is expected that product ratings positively influence users' willingness to buy.

H7b: Showing product ratings positively affects users' willingness to buy

Concerning the effect of reference prices on willingness to buy, the hypothesis is based on the reference price framework. This framework is built upon two theories: the adaptation-level theory (Helson, 1964) and the assimilation-contrast theory (Sherif & Hovland, 1961). The first theory suggests that people judge a stimulus relative to an internal reference price, which is determined by prior exposure to such stimuli. In addition to this theory, the assimilation-contrast theory suggests that based on their internal reference price (IRP), people form a distribution of prices that are considered as acceptable. In case that a stimulus (e.g. a claimed reference price) value is close to that of the internal reference price, it is assimilated with that price. On the other hand, if a stimulus is too far from the IRP, it is contrasted. Whereas internal reference prices are reference prices stored in users' memory, external reference prices can be provided to consumers through 'external' channels such as advertising and price comparison websites (Biswas & Blair, 1991).

In the field of (online) retail marketing, these external reference prices have been of particular interest. Retailers have widely adopted the concept of comparative pricing strategies, in which the actual price is compared with a higher reference price to create the perception of a discount deal (Kopalle & Lindsey-Mullikin, 2003). Prior research have found evidence of the existence of two effects of such pricing strategies. Firstly, it is suggested that external reference prices are able to increase users' perceived savings, which subsequently influences users' purchase intentions (Biswas & Blair, 1991; Grewal et al., 1998a; Lattin & Bucklin, 1989). Secondly, price discounts are likely to have a negative effect on users' perception of product quality, implying negative effects on purchase intentions (Blattberg

and Neslin, 1990). However, other studies that examined the latter effect have not found any significant evidence to support this effect, although a minimal existence of the effect has been identified (Grewal et al., 1998a; 1998b; Urbany & Bearden, 1988). Based on these findings, it is expected that showing reference prices enhances users' perception of a deal and therefore their willingness to buy.

H8b: Showing reference prices positively affects users' willingness to buy

2.4.4 Effect of perceived usability on willingness to buy

In the previous sections, the expected separate effect of the eight variables on both perceived usability and willingness to buy is discussed and predicted, based on prior studies. These hypotheses test the direct effect of these variables on PU and WtB. In almost of all the hypotheses, a positive effect is predicted. In the case of the effect of the eight variables on perceived usability, this is based on a broad amount of prior research. However, in the case of the effect on willingness to buy, the amount of studies that has examined these direct effects are limited. Often, studies that has focused on the effect of design aesthetics on purchase intentions do not examine the direct effect of design aesthetics on purchase intentions. For example, in the study of Hall and Hanna (2004) the use of preferred colors led to higher ratings of aesthetic quality, which on its turn positively influenced purchase intentions. Although it is expected that separate effects of the variables on willingness to buy exist, it is interesting to also examine the mediating effect of perceived usability. Based on the literature review in 2.4.1, 2.4.2 and 2.4.3, it is expected that a higher perceived usability leads to higher purchase intentions.

H9: Perceived usability positively influences willingness to buy

3. Methodology

Now that the theoretical and conceptual framework have been discussed and the hypotheses have been developed, the obtained knowledge should be applied in order to test the hypotheses. This chapter provides the methodology that is needed to conduct the experiment that enables testing of the various hypotheses. In the first part of this chapter, interface design of the largest Japanese online retailer application is compared with the design of the largest US online retailer application in order to derive the eight variables that serve as base of the experiment used in this thesis. Next, the actual research design is discussed, along with internal and external validity. In the third part, the measures of both the independent and dependent variables are discussed. In the fourth and last part, information is provided about the data collection, with a discussion of the survey design and a sample analysis of the respondents.

3.1 Analysis of Asian and western (US) mobile interface designs

In the introduction of this thesis (1.1), it has been pointed out that significant differences exist in m-commerce conversion rates between Asian and western countries. Although some part of these differences can possibly be explained by cultural differences, there are indications that issues with usability, technology, privacy and security are more likely to explain these differences. To find an answer on the main research question, which may on its turn explain those differences, the first step is to compare differences in mobile interface design between the two cultures. Based on this comparison, variables are derived that are used to design the experiment and test the hypotheses.

The western company that is chosen is Amazon.com, which might not come as a big surprise as they are by far the biggest online retailer in the western world (Ecommerce news, 2015; Zaczkiewicz, 2016). On a global scale though, Amazon faces fierce competition from two e-commerce powerhouses from the east: China based company Alibaba and Japan based company Rakuten (Chen, 2016). Both Asian e-commerce companies have made a large amount of acquisitions lately in order to penetrate western markets. Partially because of these acquisitions, both companies have grown substantially in the past years and it is expected that Amazon, Alibaba and Rakuten will battle for becoming the leading e-commerce company worldwide (Chen, 2016). One major difference between Rakuten and the other two companies is that Rakuten is membership-based and more loyalty-driven, focusing on creating the optimal personal shopping experience (Akhtar, 2013). Therefore, it has been decided to compare the design of Rakuten with that of Amazon, as it is expected that these differ the most of each other.

In order to find any notable differences between the designs of Amazon and Rakuten, two screenshots are analyzed to screen for differences in certain variables of mobile interface design. A graphical comparison and analysis of the screenshots can be found in figure 2 of

Appendix A. When putting both screens next to each other for comparison, one can immediately notice that both companies have designed their applications quite differently. The main differences are outlined in the table below. Main differences can be found in color design, screen density and product information. The use of the warm, red color in the Rakuten design is consistent with the findings of Cyr (2008), who state that red color is associated with happiness in more collectivist cultures. The screen density in both screenshots is also significantly different. Whereas Rakuten features just two products on the screen and only provides the name and the price of the product, Amazon features 3 products on a screen, along with an extensive amount of product information. This contains information about the product name, (reference) price, shipping costs, product ratings and other additional information. Combined with the larger number of products featured on a page, this leads to a relative high screen density.

Other, smaller differences that can be noticed are differences in navigation design, professionalism, symmetry, font design and the use of in-app advertisements. The differences as described in the table below served as base for the conceptual framework of this thesis. However, it has to be noted that difference 8, the use of advertisements, is not incorporated in the framework and the experiment. Studies has shown that (especially interruptive) ads are perceived as intrusive and have a negative influence on the willingness to buy (Acquisti & Spiekermann, 2011; McCoy et al., 2007; McCoy, Everard & Loiacono, 2009). However, it is almost impossible to test interruptive ads in the kind of experiment that is conducted in this thesis, which is a static experiment. For future research however, it could be interesting to incorporate this effect in more interactive experiments.

Table 1: Differences in screen design between Rakuten and Amazon

	Design	
	Rakuten	**Amazon**
Difference		
1 - Navigation Design	**Top**: menu/info buttons, 2 bars to search products **Bottom**: navigation bar	**Top**: menu/shopping chart buttons, 1 search bar, 1 informational/filter bar
2 - Color Design	Expressive (mostly red) colors	Neutral (mostly light grey) colors
3 - Professionalism	No lines to separate products and bars, playful shapes and icons	Lines to separate products and bars, use of rectangular shapes, more formal icons
4 - Symmetry	Symmetric product display	Asymmetric product display
5 - Screen Density	2 products featured on screen, limited use of text	3 products featured on screen, relatively large amount of text
6 - Font Design	Some variety in color and size	Large varierty in color and size
7 - Product Information	Small amount of product information, no reference pricing, no product ratings	Large amount of product information, use of reference pricing and product ratings
8 - Use of advertisement	In-app advertisement (promotional)	No in-app advertisment

3.2 Research Design

The main objectives of this research are:

- to find the importance of the separate effects of eight independent variables on two dependent variables: perceived usability and willingness to buy
- to find a possible mediating effect of the dependent variable perceived usability
- to find the optimal combination of attributes in order to optimize the mobile interface design

In order to achieve these objectives, a research design needs to be developed that is able to capture the relative importance of the various separate variables. Ultimately this should be used to develop the optimal mobile interface design, based on design aesthetics. However, there are two problems being faced in this process. The first problem is the large number of variables in this study. Although the independent variables are all dichotomous variables, the number of possible screen combinations is huge. To ensure a high internal validity of the study, a full factorial design would be the most optimal design. However, a full factorial design in this case would require $2^8 = 256$ runs to test every possible combination. This would go beyond the scope of this study. Therefore, a resolution IV fractional factorial design with a screening objective is used instead. A screening design is most commonly used to find a few significant factors from a list of many potential ones (Lawson, 2003). Typically, the number of runs in these experiments is low, because the main goal is to screen for significant factors. The fractional factorial design makes it possible to find the significance of different factors using a limited number of runs, in this case 16 runs. The design of the experiment is balanced and orthogonally designed. The complete one-sixteenth fractional factorial design, including the labels of the two-level factors, can be found in table 1 of Appendix A.

The second problem occurs with the collection of the data for this study. Ideally one would like to examine the responses of the respondents in an environment that is similar to scenario's in real life. This would guarantee a high external validity of the study. However, in order to make the experiment in this thesis similar to real life scenario's, it is needed to develop an application with functionalities similar to that of applications of mobile retailers. As this will obviously take a considerable amount of time and money to develop, this goes beyond the scope of this study. Therefore, in this study only the subjective responses are examined, since the objective responses are impossible to capture. These subjective responses will be examined by an online distributed survey, which will be further discussed in section 3.4. In order to increase the external validity of the experiment, the 16 screenshots used in the survey are based on real screenshots, with the use of existing products. Furthermore, the screenshots are about the same size as an average smartphone screen.

To increase the internal validity of the survey, the following steps are taken:

- Randomization of the 16 screens: The 16 screens presented to the respondents are presented in a random order to avoid anchoring effects.
- Constant highest/lowest price: Since there is a possibility that the willingness to buy of respondents relies on the prices of the products shown in the screens, the highest and lowest price of a product are kept constant throughout the screens. The highest priced product in every screen is priced at €49,99, whereas the lowest priced product is priced at €22,99.
- Choice of brand: Although the screens in the experiment are based on screenshots of the applications of Rakuten and Amazon, the brand that is used in the screens is Rakuten. Using screens with different brands could lead to biases based on brand preferences. Although it could be interesting to measure the effect of brands, it is not the intention of this study to do so and therefore it is preferable to use just one brand. Next to that, it is assumed that Amazon is a well known brand among the Dutch respondents, whereas Rakuten isn't. As this experiment attempts to find unbiased response effects, it is decided to use Rakuten as brand. The respondents don't recognize that brand, so it is assumed that they don't have any positive or negative feelings towards it. In contrast to Rakuten, it is assumed that respondents know Amazon as brand and therefore have certain feelings towards this brand. Using such a brand could lead to halo effects, which would decrease the internal validity of the data.
- Choice of product: The products used in this experiment are food processors. This kind of product is assumed to be a rather neutral product, implying that the respondents do not have any specific feelings towards the product. Furthermore, it is assumed that there is no considerable difference between male and female in terms of feelings towards food processors.
- Control variables: Next to the eight independent screen variables, several control variables are added in the experiment to test for possible confounding effects. Furthermore, the control variables can be used to test some of the assumptions that are discussed above (e.g. the assumption that there is no difference in gender when it comes to feelings towards food processors).

3.3 Measures and Manipulations

To measure the effect of design aesthetics on perceived usability and willingness to buy, respondents have to rate 16 different screenshots in terms of usability and willingness to buy. The screens are designed in such a way that they contain a combination of eight independent variables/screen attributes. Every screen contains a different combination of attributes.

3.3.1 Independent Variables and Manipulations

1. *Screen Density*: The value of this attribute is either Low (=0) or High (=1). Screen Density in this study refers to the number of products shown in the screen. When the screen density is low, only two products are shown in the screen. In the case of high screen density, the number of products depends on the Symmetry variable. Symmetric screens with high density contain four products, whereas asymmetric screens with high density contain three products. The difference between these two scenarios can simply be explained by the fact that it is impossible to fit four products in an asymmetric screen without considerably decreasing the readability.
2. *Navigation Design*: The value of this attribute is either Simple (=0) or Advanced (=1). A simple navigation design contains one menu bar and one search bar at the top of the screen and one navigation bar at the bottom of the screen. In contrary, screens with an advanced navigation design contain an extra search bar and an informational bar which also provides filter options.
3. *Symmetry*: The value of this attribute is either Asymmetrical (=0) or Symmetrical (=1). The definition of symmetry in this case refers to whether the products in the screen are placed symmetrical or not.
4. *Color Design*: The value of this attribute is either Neutral (=0) or Cool (=1). The color that is used in screens with neutral design is light grey, whereas the color used in screens with cool color design is blue. The changes in color occur in the background (e.g. navigation bars) and in the icons.
5. *Font Design*: The value of this attribute is either No Variety (=0) or Variety (=0). In the case that there is no variety, the font design will consist of a same size, black font. On the other hand, screens with variety in font design will contain a font with differences in size and color. Next to that, the font type will be bold.
6. *Professionalism*: The value of this attribute is either Playful (=0) or Professional (=1). Professionalism in this study refers to whether the design of the screen has a professional appearance. There will be two extreme ends in terms of professionalism. On the one end, the site looks professionally, neatly designed with a relatively high use of lines to separate products and navigation buttons and the use of rather formal icons. On the other end, the design is creative and playful with a relatively low use of lines and rectangular shapes and the use of more creative icons.
7. *Product Rating*: The value of this attribute is either No Product Rating (=0) or Product Rating (=1). This attribute is pretty straightforward; products in the screen will either contain product rating or no product rating. To avoid preferences based on the value of the product rating, products have a consistent product rating of 4,5 in the case that screens contain product ratings.
8. *Reference Price*: The value of this attribute is either No Reference Price (=0) or Reference Price (=1). The manipulation of this attribute is also straightforward. Products in the screens either contain both normal and reference prices or only normal prices.

3.3.2 Dependent Variables and Measurement

This study tests the response of respondents in terms of two dependent variables: perceived usability and willingness to buy. In prior studies, perceived usability is often measured by combining multiple items in one construct in order to test multiple dimensions of the response variable (Bruner & Kumar, 2005; Cyr et al., 2006; Koufaris, 2002). However, in this thesis there are two problems that are being faced with the data collection. Firstly, the surveys are distributed online and respondents are usually too busy to complete a survey that takes much time. Secondly, the respondents are already asked to rate 16 different screens. If the respondents have to answer multiple questions per variable per screen, it would lead to a high cognitive load. This can cause large drop-out rates or questions at the end of the survey being answered without paying sufficient attention to the screens. Waltz, Strickland and Lenz (1991) suggested that in cases like this, when respondents are busy, single item measures could be a useful alternative. Prior studies have shown that single-item scales are reliable measurement methods to measure mood and emotional states, such as job satisfaction (Wanous et al., 1997) and attitudes towards advertisements or brands (Bergkvist & Rossiter, 2007). Since perceived usability and willingness to buy can be defined as emotional responses/attitude towards screen design, it is assumed that a single-item measurement scale is sufficient to capture the responses. In the survey, the two dependent variables are measured by statements. For every of the 16 screens, respondents have to indicate whether they agree with a statement on a 0-100 scale, with 0 for total disagreement and 100 for total agreement. For the complete survey, see Appendix F.

3.4 Sample Analysis

3.4.1 Data Collection and Preparation

3.4.1.1 Collection

Due to a limited amount of time to collect a sufficient quantity of data, combined with the fact that in an experiment which examines the subjective ratings of respondents no interactivity is needed, the data of this thesis has been collected through an online survey. The survey was created with the online survey platform Qualtrics. An anonymous direct link to the survey has been distributed via social media platforms (mainly Facebook) in order to collect a sufficient amount of respondents. The survey was placed online for exactly one week, from the 18th of July until the 25th of July. In this timeframe, 96 responses have been collected.

3.4.1.2 Preparation

Out of the 96 respondents who started the survey, 76 respondents have finished the survey, implying a drop-out rate of 21%. The majority of the drop-outs occurred at the beginning of part 2, which is the part in which the respondents rate 16 screens on usability and

willingness to buy. Although some of the respondents did rate some of the screens, only 2 of the unfinished surveys had no missing values for any question. These two respondents have dropped out at the closing question, in which they could submit their mail address in order to qualify for winning the 25 euro voucher. Hence, these results were accounted as finished, useful surveys. Next, the remaining 78 surveys were investigated on missing values. This survey contained a total of 42 questions. In regard to missing values, ten percent was set as the maximum amount of missing values. This implies that 4 or less missing values were deemed as acceptable.

Out of the 78 remaining surveys, only one survey had more than 4 missing values. The data of this respondent had 16 missing values and was therefore deleted from the database. Five other surveys contained missing values (2, 2, 1, 1, 1), with a total of 7 missing values. These were all values that occurred in part 2 of the survey, implying that the respondents had not given a rating for that screen. Since these missing values contained ratings, the missing values were replaced by the mean value of the corresponding rating questions. The last step of the data preparation was to look for values that deviated from the value that was expected given the overall answers of a respondents. In total, nine values were identified that were not in line with other values given by the respondent. In all cases, this value was 0, which is probably caused by the fact that the respondent forgot to answer this question or that the sliding mechanism of the question didn't work as it was supposed to. Therefore, these values were also replaced by the mean value of the corresponding questions.

3.4.2 Participants analysis

After the recoding and cleaning of the data, a database remained containing 77 respondents. This section provides an analysis of the demographic and behavioral characteristics of the participants in the sample. Below, some of the key characteristics are summarized. Graphical presentation of these characteristics can be found in figure 3 of Appendix A.

- The majority of the 77 respondents is female (61% vs. 39% male, figure 3A).
- Roughly half of the respondents is currently in possession of a food processor (figure 3B).
- Almost three-quarter of the respondents is currently not interested in buying a food processor (figure 3C).
- The frequency of online purchases made via smartphones is surprisingly low, with only 9 out of 77 respondents indicating to frequently or often make purchases on their smartphone (figure 3D).
- Regarding buying a product online, price and reliability of the website are the most important aspects according to the respondents, rating over 80 on average (of a maximum rating of 100). On the contrary, product range is indicated to be of considerably lower importance, rating just under 50 on average (figure 3E).

4. Results

In this chapter, results of the conducted experiment are presented. Firstly, an analysis of the ratings of the 16 different screens is performed, including average ratings of perceived usability, willingness to buy and overall rating. Additionally, it is examined whether there are significant differences between certain groups, based on binary control variables. Secondly, various regression models are tested in order to discover the relative importance of the screen variables, to test the effect of control variables and to test for the existence of significant interaction variables. The third and last part will test the hypotheses that were developed in chapter 2.4.

4.1 Screen ratings analysis

Part two of the survey, in which the respondents are shown 16 different screens which they have to rate on both usability and willingness to buy, serves as framework for the conducted experiment. Each of the 16 screens contained a different set of combinations of the 8 variables that were included in the screen designs. In this section, an analysis of the average ratings of the different screens is performed. Ratings of Perceived Usability, Willingness to Buy, Overall Rating and differences between groups are analyzed respectively.

4.1.1 Perceived Usability

The first dependent variable that is analyzed is the Perceived Usability (PU) of the screens. Respondents were asked to indicate the user friendliness/usability of a specific screen on a scale ranging from 0 to 100. Table 2 provides the average mean ranking of the 16 different screens, sorted from high to low. The second column provides information about the average standard deviation of the screen ratings. The total average rating of PU is 64,42 (SD = 17,087). Looking at Table 2, it can be concluded that screen 6 is perceived as the most usable screen (M = 70,55, SD= 14,773), being rated almost 2 points higher than the second highest rated screen, screen 2 (M = 68,91, SD = 15,440). At the bottom of the table, two screens are clearly perceived as least usable, both scoring below a rating of 60 (Screen 11: M = 59,60, SD = 17,994; Screen 10: M = 59,55, SD = 17,936). Furthermore, an ANOVA Between Subjects analysis was performed, out of which can be concluded that the groups are significantly different from each other (p = 0,000).

Perceived Usability		
Screen	**Mean**	**SD**
6	70,55	14,773
2	68,91	15,440
16	67,88	15,584
13	67,81	15,281
15	66,42	15,302
9	66,30	17,845
8	65,06	17,688
7	64,69	16,409
12	64,13	17,325
1	62,88	16,994
3	62,42	17,428
4	61,73	19,187
5	61,71	18,413
14	61,04	15,649
11	59,60	17,994
10	59,55	17,936
Total	**64,42**	**17,087**
ANOVA Between Subjects: F-value: 3,038*		
1) * significant at 1%-level		

Table 2: Ranking of PU Ratings

4.1.2 Willingness to buy

Table 3: Ranking of WtB Ratings

Willingness to buy		
Screen	**Mean**	**SD**
6	59,09	19,507
13	58,30	17,840
16	57,75	18,793
2	56,40	19,444
9	55,30	20,146
15	55,23	17,352
7	54,40	18,808
12	53,99	18,854
5	51,83	19,969
3	51,71	19,378
8	51,31	17,535
1	50,91	18,471
4	50,83	22,069
14	50,64	18,312
10	49,35	19,913
11	48,22	19,166
Total	**53,45**	**19,284**
ANOVA Between Subjects: F-value: 2,328*		
1) * significant at 1%-level		

The average ratings of the second dependent variable, willingness to buy, are shown in Table 3. To test this variable, respondents were asked to indicate whether the screen design motivated them to buy a food processor (on a scale ranging from 0 to 100). One thing that can immediately be noted is that the average ratings of WtB are considerably lower than the average ratings of perceived usability (total average: M = 53,45, SD = 19.284). Furthermore, it can be noted that screen 6 also tops the list for willingness to buy (M = 59,09, SD = 19,507), followed by screen 13 (M = 58,30, SD = 17,840) which ranked fourth on perceived usability. At the bottom of the table, another similarity with the PU ratings can be identified as screen 10 and 11 close the rankings again, although they switched places this time. Both screens ranked below a rating of 50 (M = 49,35, SD = 19,913; M = 48,22, SD = 19,166, respectively). In this case, the ANOVA Between Subjects test also suggests that the 16 screens significantly differ from each other (p = 0,003).

4.1.3 Overall screen ratings

Next to the ratings of the separate dependent variables, it is interesting to get insights of a combination of both variables. In order to get these insights, ratings of the 16 screens were combined to create an overall screen rating. Screen 6, which was rated as top screen in both perceived usability and willingness to buy, is rated as overall top screen with an average overall mean of 64,82. The top 3 screens is completed by screen 13 (M = 63,05) and screen 16 (M = 62,82). At the bottom of the rank, screen 10 and 11 are rated over 10 points lower on average than the top rated screen (M = 54,45 and 53,91, respectively). Table 4 shows the complete ranking of overall screen ratings, with an average overall screen rating of 58,94.

In order to get a visual view of the overall screen ratings, the 16 screens are sorted from high to low based on overall screen rating (Appendix B). The visual presentation of the ranking could already provide some insights of the relative importance of the attributes. Looking at the top 4 rated screens, some consistent similarities can be identified. Firstly, it can be noted that the color of the top 4 rated screens is neutral. Moreover, when looking at the bottom 6 screens, it can be noted that all of them include the blue color. This raises the suspicion that neutral colors are perceived as more usable and lead to an increased willingness to buy.

The second screen variable that is consistent in all four top rated screens is the variable navigation design. In all four screens, the navigation design is considered to be an advanced navigation design, with two search bars and one informational/filter bar at the top of the screen and one navigational bar at the bottom of the screen.

These findings may suggest that respondents perceive additional navigation options as more usable, which makes them more likely to buy a product. However, having a closer look at the four lowest rated screens, this suspicion may have to be revised because three out of the four lowest rated screens also include an advanced navigation design. This inconsistency disputes the earlier suspicion of the importance of navigation design and suggests that the presence of an advanced navigation design could rather be coincidental, caused by the effect of other, more significantly important variables.

Table 4: Overall Ratings

Overall Rating			
Screen	**Mean PU**	**Mean WtB**	**Overall Mean**
6	70,55	59,09	64,82
13	67,81	58,30	63,05
16	67,88	57,75	62,82
2	68,91	56,40	62,65
15	66,42	55,23	60,82
9	66,30	55,30	60,80
7	64,69	54,40	59,54
12	64,13	53,99	59,06
8	65,06	51,31	58,19
3	62,42	51,71	57,06
1	62,88	50,91	56,90
5	61,71	51,83	56,77
4	61,73	50,83	56,28
14	61,04	50,64	55,84
10	59,55	49,35	54,45
11	59,60	48,22	53,91
Total	**64,42**	**53,45**	**58,94**

The last screen variable that is consistent in all four top rated screens is the product rating variable since all four screens contain product ratings. Moreover, the presence of product rating in the screens is even consistent at the top six rated screens. Comparing this with the bottom four rated screens, only one of those four screens contains product rating. Although for now it can't be concluded that there is a significant effect of product rating on PU and WtB, there are signs that such an effect exists. To test for significant effects of this and the other seven attributes, various regression analyses will be conducted.

4.1.4 Between Group Differences

Now that the screen ratings have been ranked and analyzed, it could be interesting to test for significant differences between various groups in terms of perceived usability and willingness to buy. In this section, these tests will be performed for three control variables, which are based on questions asked in part one of the survey. These control variables are: gender, food processor ownership and interest in buying a food processor. All three variables are binary variables and are coded as dummy variables. Significant differences between groups can be an indicator for significant influences of specific variables in regression models, although this isn't necessarily true. Since the three control variables are binary variables, the dependent variables (PU and WtB) are continuous variables and the groups that are tested are unrelated to each other, the differences between the groups can

be tested using an independent sample t-test. The results of the t-tests on the three different control variables are shown in the table below:

Table 5: Differences Between Groups

	Gender	Mean	Difference	FPO[1]	Mean	Difference	IBFP[2]	Mean	Difference
Perceived Usability	Female	64,52	0,26	Yes	66,28	3,68*	Yes	63,42	1,37
	Male	64,26		No	62,60		No	64,79	
Willingness to Buy	Female	55,88	6,23**	Yes	56,14	5,30**	Yes	58,15	6,46*
	Male	49,65		No	50,84		No	51,69	

1) [1] = Food Processor Ownership

2) [2] = Interest in Buying a Food Processor

3) * = statistically significant at 1%-level, no violation of equal variances assumption

4) ** = statistically significant at 1%-level, violation of equal variances assumption

The first control variable that is tested is gender. As can be concluded from Table 5, there is no significant difference between ratings of perceived usability between male and female ($p = 0,782$). However, the difference in ratings in terms of willingness to buy is rather large (6,23). It is therefore not surprising that this difference is found to be significant at a 1%-level ($p = 0,000$). Furthermore, it has to be noted that for both PU and WtB, the assumption of equal variances was violated (Levene's Test's $p < 0,05$ in both cases). Because of this violation, the Welch's t-test was used, which assumes that variances are not equal.

The second binary control variable that is tested is whether respondents are in possession of a food processor (0 = no, 1 = yes). The conclusions that can be drawn from Table 5 is that for both perceived usability and willingness to buy, there are statistically significant differences between the respondents who own and who do not own a food processor ($p = 0,000$ in both cases). In case of PU, the assumption of equal variances was not violated (Levene's Test -> $p = 0,253$) and therefore an independent sample t-test was conducted. However, in the case of WtB, this assumption was violated (Levene's Test -> $p = 0,000$) and therefore the Welch's t-test was used to test for differences between the groups.

The last variable that is tested is whether respondents are interested in buying a food processor at the moment that the survey was distributed (0 = no, 1 = yes). As can be concluded from the figures in Table 5, in terms of perceived usability, there is no statistically significant difference between respondents who are interested and respondents who are not interested in buying a food processor ($p = 0,211$). On the other hand, there is a clear, significant differences between these two groups when it comes to willingness to buy, with a p-value that is significant at a 1%-level ($p = 0,000$). Furthermore, it has to be noted that in both cases the variances were pretty equal and therefore the assumption of equal variances was not violated (p values are 0,581 and 0,942, respectively), enabling the performance of an independent sample t-test.

4.2 Regression Models

The findings in section 4.1 already suggested that there are significant differences in average screen rating between the screens itself and differences between several groups, based on three binary control variables. Moreover, consistency of various screen variables in the top rated screens indicate that some screen variables have a significant effect on the dependent variables (ratings of PU and WtB). In order to reveal the separate effects of the screen variables, the effect of the control variables and the existence of possible interaction variables, this section uses various regression models in order to examine these effects. Subsequently, the findings of these models are used to test the hypotheses that were stated in section 2.4. The regression technique used in this thesis is a multiple regression (MR), which examines the relationship between a single outcome measure and several predictor or independent variables. However, in order to perform a complete regression analysis, it is critical that several assumptions are met (Osborne & Waters, 2002). When these assumptions are not met, there is a chance that the significance of the effect size is over- or under-estimated. Therefore, it is necessary to check for any violation of these assumptions.

In Appendix E, these four assumptions are tested, along with tests of normality of the two dependent variables. Based on the results on this appendix, it can be concluded that all assumptions of a multiple regression model are met and therefore the use of a multiple regression is justified. Next, various multiple regression models are applied to both dependent variables in order to reveal the relationships between different variables and to test for significant effects.

4.2.1 Multiple Regression Models applied to Perceived Usability

The variables in this thesis can roughly be divided into three different categories: dependent rating variables, independent screen variables and independent control variables. The main objective of this research is to examine the separate effect of the screen attributes on both PU and WtB. This results in the first multiple regression model, which only tests the effect of the screen variables. Control variables are not incorporated in this model. This model, which is from now on denoted as the basic model, looks as follows for the dependent variable PU, in which screen variables are noted as SV:

$$PU_i = \beta_0 + \beta_1 SV_{1i} + \beta_2 SV_{2i} + \beta_3 SV_{3i} + \beta_4 SV_{4i} + \beta_5 SV_{5i} + \beta_6 SV_{6i} + \beta_7 SV_{7i} + \beta_8 SV_{8i} + \varepsilon_i$$

Running this model in SPSS gives the output shown in table 6 on the next page. As can be concluded from this output, the constant factor is 63,769 and is highly significant. Furthermore, three of the eight screen variables have a significant effect on Perceived Usability: Screen Density, Color Design and Product Rating. The effects of these variables are all statistically significant at a 1%-level. The other five screen variables doesn't seem to have a large influence on PU, although Symmetry seems to have some (non-significant, p = 0,151) influence. Next to that, the R^2 of the basic model suggests that this model explains merely 4,1 percent of the response variable variation. The last conclusion that can be drawn from

the basic model output is that the model has significantly more predictive value than the intercept-only model (ANOVA F statistic is sign. at 1%-level).

Table 6: Basic Model PU

Coefficients[a]

Basic Model**		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	63,769	1,388		45,931	,000*
	Screen Density	2,793	,924	,085	3,022	,003*
	Navigation Design	-,005	,924	,000	-,005	,996
	Symmetry	-1,327	,924	-,040	-1,435	,151
	Color Design	-4,228	,924	-,129	-4,574	,000*
	Font Design	,508	,924	,015	,550	,583
	Professionalism	1,042	,924	,032	1,127	,260
	Product Rating	3,912	,924	,119	4,233	,000*
	Reference Price	-,460	,924	-,014	-,497	,619

a. Dependent Variable: Perceived Usability

1) * = Statistically significant at 1%-level

2) ** = R Square of model: 0,041, ANOVA: F: 6,510 -> p = 0,000*

The basic model only takes into account the effect of the eight screen variables on PU. However, since usability is a rather broad concept, it is reasonable to assume that there are many other factors that influence this variable. This is confirmed by the low R^2 of the basic model. Therefore, it is interesting to add control variables, which were derived from the first part of the survey. The resulting model, let's call it the Extended Model, will look as follows for the dependent variable PU, in which SV are screen variables and CV are control variables:

$$PU_i = \beta_0 + \beta_1 SV_{1i} + \beta_2 SV_{2i} + \beta_3 SV_{3i} + \beta_4 SV_{4i} + \beta_5 SV_{5i} + \beta_6 SV_{6i} + \beta_7 SV_{7i} + \beta_8 SV_{8i} + \beta_9 CV_{1i} + \beta_{10} CV_{2i} + \beta_{11} CV_{3i} + \beta_{12} CV_{4i} + \beta_{13} CV_{5i} + \beta_{14} CV_{6i} + \beta_{15} CV_{7i} + \beta_{16} CV_{8i} + \beta_{17} CV_{9i} + \beta_{18} CV_{10i} + \varepsilon_i$$

The output of the extended multiple regression model can be found in Appendix C, table 2. For the sake of readability, only variables with a p-value of 0,15 or lower are included in table 7 on the next page. When looking at the output of the extended model, a few things can be noted. Firstly, the overall significance of the screen variables has increased. Screen Density, Color Design and Product Rating still have significant effects at a 1%-level after the addition of control variables. Focusing on the effects of the control variables, it can be concluded that only two of the ten control variables have a significant effect on PU. However, three other control variables show marginally significant effects at a 10%-level, whereas the variable interest in buying a food processor has a p-value just outside of the 10%-level. To conclude the analysis of the extended model for PU, it can be noted that the R^2 has more than doubled, now explaining 9,7 percent of the response variable variation.

Table 7: Extended Model PU (Adjusted output)

Coefficients[a]

Extended Model***	Unst. B	t	Sig.
(Constant)	45,301	9,366	,000*
Screen Density	2,805	3,115	,002*
Symmetry	-1,335	-1,482	,139
Color Design	-4,244	-4,713	,000*
Product Rating	3,899	4,329	,000*
Frequency of buying online	1,023	2,153	,031**
Impact of Price	,072	1,655	,098
Impact of Reviews	,049	1,714	,087
Impact of Product Range	,141	4,623	,000*
Food Processor Ownership	1,842	1,928	,054
Interest in buying a Food Processor	-1,664	-1,578	,115

a. Dependent Variable: Perceived Usability
1) * = Statistically significant at 1%-level
2) ** = Statistically significant at 5%-level
3) *** = R Square of model: 0,097, ANOVA: F: 7,197 -> p = 0,000*

The extended model tested the separate effect of each screen- and control-variable. However, the model has not tested interaction effects. Adding interaction terms can considerably increase the understanding the relationships between certain variables and increase the R^2 of the model. In the extended model, it is interesting to see if there are any significant interaction variables that increase the understanding of some relationships. One of those interaction terms is included in the hypotheses (H3c), which expects a difference between male and female in terms of the effect of symmetry on PU. To test this hypothesis, interaction variables have to be computed in SPSS.

Computation of this variable is performed by multiplying the variables Symmetry and Gender. Since both Symmetry and Gender are dummy variables, this will basically give a value of 1 for cases in which the products in the screen are symmetrically placed and the respondent is male. In any other case, the value will be 0. Next to this, it could be interesting to see if there are any other significant interaction variables based on gender, or at least have a marginal effect. Therefore, the variable Gender is multiplied by every other dummy variable in the model, including food processor ownership and interest in buying a food processor.

Next, these variables are incorporated in the regression model to check for any significant effects. For the other two dummy variables, food processor ownership and interest in buying a food processor, the same procedure is followed. An example of this procedure, in which the dummy variables are multiplied by gender, can be found in Appendix C, table 3. In the case of table 3, there are two interaction variables (MaleCD and MaleFreqBO) that seem to have at least a marginal effect on PU ($p < 0,15$). Therefore, these two variables will be

included in the full model. Repeating of this procedure leads to the following interaction variables to be included in the full model: MaleSym, MaleCD, OwnFP_BuyFP, Int_Sym and Int_CD. The corresponding Beta's and significance levels can be found in table 5 of appendix C.

As can be concluded from that table, a total of five interaction variables is deemed to have at least a marginal effect on perceived usability, with the exception of MaleSym that is included anyway because this variable must be used to test hypothesis 3c. The five interaction variables from table 8 will be added to the extended model in order to form the full multiple regression model. This model will looks as follows, in which the interaction variables are noted as IV:

$$PU_i = \beta_0 + \beta_1 SV_{1i} + \beta_2 SV_{2i} + \beta_3 SV_{3i} + \beta_4 SV_{4i} + \beta_5 SV_{5i} + \beta_6 SV_{6i} + \beta_7 SV_{7i} + \beta_8 SV_{8i} + \beta_9 CV_{1i} + \beta_{10} CV_{2i} + \beta_{11} CV_{3i} + \beta_{12} CV_{4i} + \beta_{13} CV_{5i} + \beta_{14} CV_{6i} + \beta_{15} CV_{7i} + \beta_{16} CV_{8i} + \beta_{17} CV_{9i} + \beta_{18} CV_{10i} + \beta_{19} IV_{1i} + \beta_{20} IV_{2i} + \beta_{21} IV_{3i} + \beta_{22} IV_{4i} + \beta_{23} IV_{5i} + \varepsilon_i$$

The interaction variables itself are officially noted somewhat differently, including the interaction term. For example MaleCD would be noted as $\beta_{20} SV_{4i} CV_{1i}$. The full model is also run on SPSS, which leads to the output of table 7 in Appendix C. The adjusted output of the full model, with only variables with a $p < 0,15$ incorporated in the table, can be found below:

Table 8: Full Model PU (Adjusted output)

Coefficients[a]

Full Model***	Unst. B	St. B	t	Sig.
(Constant)	45,797	4,672	9,803	,000*
Screen Density	2,841	,861	3,298	,001*
Symmetry	-2,211	1,247	-1,773	,076
Color Design	-4,761	1,247	-3,818	,000*
Product Rating	3,854	,861	4,474	,000*
Frequency of buying online	,918	,455	2,018	,044**
Impact of Price	,099	,042	2,355	,019**
Impact of Reviews	,055	,027	2,022	,043**
Impact of Reliability	-,085	,031	-2,696	,007*
Impact of Product Range	,153	,029	5,240	,000*
Food Processor Ownership	7,816	1,079	7,244	,000*
Interest in buying a Food Processor	7,369	1,919	3,840	,000*
MaleCD	3,101	1,767	1,755	,080
OwnFP_BuyFP	-21,504	2,063	-10,424	,000*
Int_Sym	3,057	1,942	1,574	,116
Int_CD	-2,762	1,942	-1,422	,155

a. Dependent Variable: Perceived Usability
1) * = Statistically significant at 1%-level
2) ** = Statistically significant at 5%-level
3) *** = R Square of model: 0,178, ANOVA: F: 11,215 -> p = 0,000*

One thing that can immediately be noted when looking at the output of table 8 is the fact that in the full model, there are considerably more variables that have a significant effect on perceived usability. A total of eleven variables (3 screen variables, 7 control variables and 1 interaction variable) have statistically significant effects at a 5%-level, with 8 of those 11 effects being also significant at a 1%-level. Furthermore, it can be noted that the screen variable Symmetry has considerably increased in significance, although it is still not significant at a 5%-level. The interaction MaleCD also seems to have a marginal significant effect on PU, with a p-level that lies within the 10%-level (p = 0,080). Lastly, the R^2 of the full model is almost twice as high as the R^2 of the extended model (0,178 vs. 0,097), now being able to explain 17,8 percent of the response variable variation.

4.2.2 Multiple Regression Models applied to Willingness to Buy

In order to test for significant relationships between independent variables and willingness to buy, the same procedure can be followed as for perceived usability. This implies that the three multiple regression models of the previous section can also be applied to WtB. The basic model and the extended model will look pretty much exactly the same, the only difference is that PU is replaced by WtB:

Basic Model:

$$WtB_i = \beta_0 + \beta_1 SV_{1i} + \beta_2 SV_{2i} + \beta_3 SV_{3i} + \beta_4 SV_{4i} + \beta_5 SV_{5i} + \beta_6 SV_{6i} + \beta_7 SV_{7i} + \beta_8 SV_{8i} + \varepsilon_i$$

Extended Model:

$$WtB_i = \beta_0 + \beta_1 SV_{1i} + \beta_2 SV_{2i} + \beta_3 SV_{3i} + \beta_4 SV_{4i} + \beta_5 SV_{5i} + \beta_6 SV_{6i} + \beta_7 SV_{7i} + \beta_8 SV_{8i} + \beta_9 CV_{1i} + \beta_{10} CV_{2i} + \beta_{11} CV_{3i} + \beta_{12} CV_{4i} + \beta_{13} CV_{5i} + \beta_{14} CV_{6i} + \beta_{15} CV_{7i} + \beta_{16} CV_{8i} + \beta_{17} CV_{9i} + \beta_{18} CV_{10i} + \varepsilon_i$$

The Basic Model and the Extended model are combined in one model using a Hierarchical Multiple Regression. This type of MR basically shows the first and second model in one table. This can be useful to easily compare the effects and significance of the independent variables of both models. The output of the Hierarchical Model can be found in table 8 of Appendix C. The first model represent the Basic Model, the second model is the Extended Model. An adjusted output of the Hierarchical Model, which only includes independent variables with a p-value of <0,15, is shown in Table 9 on the next page.

Analysis of the output in Table 9 shows that in model 1, the Basic Model, only two of the eight screen variables (Color Design and Product Rating) have significant effects on WtB, both being significant at a 1%-level. When adding the control variables, which leads to model 2, it can be noted that 7 out of the total of 10 control variables have significant effects on WtB. Moreover, the dummy control variables seem to have very strong influence on WtB, considering their high t-values and significance at a 1%-level. The last thing that can be concluded is that the R^2 of model 2 is considerably higher than the R^2 of the Basic Model (0,140 vs. 0,027), explaining a total of 14 percent of the response variable variation.

Table 9: Hierarchical Model WtB (Adjusted output)

Coefficients[a]

Hierarchical Model		Unst. B	St. B	t	Sig.
1***	(Constant)	54,483		32,675	,000*
	Screen Density	1,646	,043	1,513	,130
	Color Design	-3,831	-,099	-3,522	,000*
	Product Rating	4,276	,111	3,931	,000*
2****	(Constant)	44,933		8,144	,000*
	Screen Density	1,646	,043	1,603	,109
	Color Design	-3,831	-,099	-3,732	,000*
	Product Rating	4,276	,111	4,165	,000*
	Gender	-4,750	-,120	-4,265	,000*
	Impact of Ease to find products	,094	,098	3,042	,002*
	Impact of Reliability	-,123	-,106	-3,363	,001*
	Impact of Product Range	,072	,080	2,077	,038**
	Impact of Product Info	,130	,112	2,984	,003*
	Food Processor Ownership	4,625	,120	4,244	,000*
	Interest in buying a Food Processor	6,768	,156	5,635	,000*

a. Dependent Variable: Willingness to Buy
1) * = Statistically significant at 1%-level
2) ** = Statistically significant at 5%-level
3) *** = Basic Model: R Square: 0,027, ANOVA: F: 4,168 -> p = 0,000*
4) **** = Extended Model: R Square: 0,140, ANOVA: F: 10,952 -> p = 0,000*

Now that the separate effects of the various independent variables have been examined by using a Hierarchical Multiple Regression, it is interesting to check for any significant interaction variables. In order to achieve this, the same procedure is followed as for the other dependent variable, implying that interaction variables are computed based on the three dummy control variables. This procedure leads to the identification of the following interaction variables, which are thought to have at least a marginally significant influence on WtB: MaleCD, MaleOwnFP, OwnFP-BuyFP and Int_CD. The corresponding Beta's and significance levels can be found in table 6 of appendix C.

The addition of the four interaction variables that table leads to the full multiple regression model of willingness to buy. This model looks as follows, again with the interaction variables noted as IV:

$$WtB_i = \beta_0 + \beta_1 SV_{1i} + \beta_2 SV_{2i} + \beta_3 SV_{3i} + \beta_4 SV_{4i} + \beta_5 SV_{5i} + \beta_6 SV_{6i} + \beta_7 SV_{7i} + \beta_8 SV_{8i} + \beta_9 CV_{1i} + \beta_{10} CV_{2i} + \beta_{11} CV_{3i} + \beta_{12} CV_{4i} + \beta_{13} CV_{5i} + \beta_{14} CV_{6i} + \beta_{15} CV_{7i} + \beta_{16} CV_{8i} + \beta_{17} CV_{9i} + \beta_{18} CV_{10i} + \beta_{19} IV_{1i} + \beta_{20} IV_{2i} + \beta_{21} IV_{3i} + \beta_{22} IV_{4i} + \varepsilon_i$$

Running this regression model in SPSS gives the output shown in table 9 of Appendix C. The adjusted output can be found in the table below:

Table 10: Full Model WtB (Adjusted output)

Coefficients[a]

Full Model***	Unst. B	St. B	t	Sig.
(Constant)	50,881		9,552	,000*
Screen Density	1,646	,043	1,682	,093
Color Design	-4,447	-,115	-3,145	,002*
Product Rating	4,276	,111	4,369	,000*
Gender	2,451	,062	1,402	,161
Impact of Price	-,116	-,067	-2,391	,017**
Impact of Reviews	,073	,077	2,343	,019**
Impact of Reliability	-,178	-,153	-4,993	,000*
Impact of Product Range	,142	,157	4,183	,000*
Food Processor Ownership	16,798	,436	11,038	,000*
Interest in buying a Food Processor	17,295	,400	9,132	,000*
MaleCD	3,818	,078	1,898	,058
MaleOwnFP	-19,058	-,359	-8,715	,000*
OwnFP_BuyFP	-17,197	-,272	-7,317	,000*
Int_CD	-3,194	-,057	-1,450	,147

a. Dependent Variable: Willingness to Buy
1) * = Statistically significant at 1%-level
2) ** = Statistically significant at 5%-level
3) *** = R Square of Model: 0,221, ANOVA: F: 15,588 -> p = 0,000*

Looking at the output of the full model, the following things can be concluded:

- Two of the eight screen variables have a significant effect on WtB, whereas Screen Density has become slightly more significant than in the extended model, now being marginally significant at a 10%-level.
- Six control variables have a significant influence on WtB. Moreover, the influence of Food Processor Ownership and Interest in buying a food processor is remarkably high. This was already noted within the extended model, however in the full model these effects have even tripled according to the standardized Beta.
- The control variable Gender, which had a highly significant effect in the extended model (p = 0,000), is insignificant in the full model (p = 0,161). This can be explained by the introduction of two interaction variables that include gender (MaleCD and MaleOwnFP). In particular the effect of the latter interaction variable is highly significant, which can explain a big part of why gender was significant in the extended model.
- Next to the remarkably high significant effect of MaleOwnFP, a second interaction variable also proves to have a strong, significant effect on willingness to buy. This variable, OwnFP_BuyFP, represents the case in which respondents are in possession of a food processor but still interested in buying a food processor. If this is the case, respondents rate the screens considerably lower in terms of willingness to buy.

- The R^2 of the full multiple regression model has considerably increased with the addition of the four interaction variables (0,221 vs. 0,140), which implies that the model now explains a total of 22,1 percent of the variance of willingness to buy. When compared to the full multiple regression model of perceived usability, it can be concluded that the full model of WtB explains a larger part of the variance of the response variable (0,221 vs. 0,178).

4.3 Testing of hypotheses

In the previous sections of this chapter, various statistical tests have been performed in order to reveal relationships between the independent and dependent variables and the significance of those relationships. The main interest of this study is to discover whether the eight screen attributes that are tested in this study have any significant influence on perceived usability and willingness to buy. In order to test these relationships, a set of hypotheses was developed in section 2.4. In this section, the findings of sections 4.1, 4.2 are combined to test whether these hypotheses are rejected or not. The hypotheses will be tested per group (i.e. Classical Aesthetics, Expressive Aesthetics and Product Specific Content).

4.3.1 Classical Aesthetics (CA)

This category includes the following variables: Screen Density, Navigation Design and Symmetry. In this section, it is merely concluded whether the hypotheses are accepted or rejected. Discussion of these outcomes is covered in the next chapter.

The hypotheses of the first screen variable, Screen Density, are as follows:

H1a: Interface designs with low screen density will lead to higher perceived usability
H1b: Interface designs with low screen density will lead to higher willingness to buy

In the database Screen Density was coded as a dummy variable with a value of 0 for low screen density and a value of 1 for high screen density. For the regression model, this basically means that in case of a significant negative Beta, the 0 hypothesis is accepted. Looking at the regression models of perceived usability, it can be concluded that in all three models the effect of Screen Density on PU was found to be significant at a 1%-level (p = 0,003, 0,002 and 0,001 respectively). However, when looking at the Beta of the effects, the results are not as expected. The Beta of the relationships is positive (B = 2,793, 2,805 and 2,841 respectively), which implies that low screen density leads to a lower perceived usability. Therefore, H1a is rejected. With regard to the effect of Screen Density on willingness to buy, the exact same problem with the Beta can be noted (B = 1,646 in all three models). Moreover, in neither of the three models the effect of Screen Density has a significant effect on WtB at a 5%-level, although in the full model it can be said that the variable has a marginally significant influence at a 10%-level. However, it is clear that H1b is also rejected.

Moving on to the second screen variable associated to classical aesthetics, Navigation Design, the proposed hypotheses are as follows:

H2a: Clear and simple navigation design positively affects perceived usability
H2b: Clear and simple navigation design positively affects willingness to buy

In part 1 of this chapter, overall screen ratings were calculated. Subsequently the screens were sorted in order from high to low, after which this ranking was visually examined to check for any preliminary suspicions of relationships between variables. Based on the top rated screens, there was a suspicion that navigation design could have a significant influence on the dependent variables. However, since the trend was not consistent throughout all screens this suspicion was immediately disputed.

In the database, Navigation Design was coded as a dummy variable with a value of 0 for simple design and 1 for advanced design. In order to accept hypotheses H2a and H2b, the Beta of navigation design should therefore be negative and significant. Although the Beta of navigation design is negative in all three regression models of the dependent variable PU, it is noted that the effects are found to be highly insignificant, with a p-value of over 0,9 in all models. Therefore, H2a is rejected. When looking at the regression models of WtB, the effect of navigation design is considerably more significant but still far from statistically significant ($p > 0,4$ in all models). Next to that, the Beta in these models is positive whereas it was hypothesized that there is a negative Beta. This leads to the conclusion that H2b is also rejected.

The last variable that was associated with classical aesthetics is symmetry. The hypotheses that were developed for this variable are:

H3a: Symmetry positively affects users' perceived usability
H3b: Symmetry positively affects users' willingness to buy
H3c: Aesthetic judgements of male participants regarding symmetry are significantly different than those of female participants

In the database, Symmetry was coded as a dummy variable with a value of 0 for assymetrical design and 1 for symmetrical design. This implies that in order to accept the 0 hypothesis, the Beta of the Symmetry variable should be positive and significant for both dependent variables. Looking at the regression models of PU and WtB it is interesting to see that the Beta is once again opposite of the hypothesized Beta, since for both dependent variables there is a negative Beta of Symmetry. Moreover, in all six regression models the effect of symmetry is insignificant at a 5%-level. However, in the full model of perceived usability the effect of symmetry is close to being significant at a 5%-level ($p = 0,076$). Next to that, a third hypothesis was tested for the symmetry variable which assumed that the aesthetics judgements of male respondents regarding symmetry are significantly different from those of female. The interaction variable MaleSym was computed to test for indications of this

effect. However, the p-value was highly insignificant (0,973). Combined with the other findings on symmetry, it can be concluded that all three hypotheses are rejected.

4.3.2 Expressive Aesthetics

This category includes the following three variables: Color Design, Font Design and Professionalism. For the first of these variables, the following hypotheses were developed:

H4a: The use of extreme short wavelength colors positively influences perceived usability
H4b: The use of extreme short wavelength colors positively influences willingness to buy

The visual examination of the overall average screen ratings provided some interesting preliminary findings, as there seemed to be a consistent trend of the color design in regards to overall ratings, with screens that included a neutral color design being rated higher on overall. However, this suspicion had to be analyzed more in depth with regression analyses in order to confirm the significance of this trend. When analyzing the output of the six regression models, it can be concluded that this suspicion is found to be true. In all of the six models, the variable has a negative Beta that is statistically significant at a 1%-level. However, once again the Beta was exactly the opposite of what was hypothesized. Therefore, both H4a and H4b are rejected.

Concerning the second variable associated with expressive aesthetics, Font Design, the following hypotheses were developed:

H5a: Variety in font design positively affects perceived usability
H5b: Variety in font design positively affects willingness to buy

In the database, Font Design was coded as a dummy variable with a value of 0 for no variety in font design and 1 for variety in font design. In order to accept hypotheses H5a and H5b, the Beta of font design should therefore be positive and significant. When looking at the regression models of PU, it can be noted that the Beta of Font Design is indeed positive in all three models. However, the effect of Font Design on perceived usability is insignificant with a p-value of over 0,3 in all models. Moving on to the three regression models of WtB, one comes to the conclusion that the effect of font design is even more insignificant in this case (p-value > 0,7 in all models). The fact that next to that the Beta is negative instead of positive makes even a stronger case to state that both H5a and H5b are both rejected.

The third and last variable that is associated to expressive aesthetics is professionalism. The developed hypotheses for this variable are:

H6a: Professional appearance negatively affects perceived usefulness
H6b: Professional appearance positively affects willingness to buy

Like all screen variables, this variable was also coded as dummy variable, with a value of 0 in case of a playful design and a value of 1 for a professional assign. This implies that in order to accept the first hypothesis, the Beta should be negative and significant, whereas in case of H6b the Beta should be positive and statistically significant. When looking at the regression models of perceived usability, it can once again be concluded that the Beta is in contrast with the Beta that was hypothesized, since the Beta is positive in all three models. Next to that, the effect is proven to be insignificant with a p-value ranging from 0,260 in the basic model to 0,220 in the full model. In the regression models that investigates the effect on WtB the Beta is also positive, which is in accordance with the hypothesis. However, in this case it can also be concluded that the effect is insignificant with a p-value that is larger than 0,5 in all models. Therefore, H6a and H6b are rejected.

4.3.3 Product Specific Content (PSC)

The last category is Product Specific Content. The variables associated with this category are the presence of product ratings and the presence of a reference price. Along with Color Design, the presence of product rating in screens was found to be a consistent variable in the top rated screens in the visual analysis of the overall screen ratings ranking. This raised suspicion that this variable could have a significant influence on both PU and WtB. This preliminary findings would be in accordance with the developed hypotheses for Product Rating, which are:

H7a: Showing product ratings increases users' perceived usability
H7b: Showing product ratings positively affects users' willingness to buy

In order to test the presence and significance of the effects hypothesized above, the regression models of PU and WtB are analysed again. In case that the 0 hypotheses are accepted, the Beta of Product Rating needs to be positive and significant. It requires only a quick glimpse at the models to draw a conclusion on this matter. In all six regression models, the Beta is positive and highly significant at a 1%-level. Therefore, both H7a and H7b are accepted.

The last variable that is tested is the presence of (external) reference prices. The corresponding hypotheses are:

H8a: Showing an external reference price increases users' perceived usability
H8b: Showing reference prices positively affects users' willingness to buy

Considering the fact that Reference Price is also a dummy variable, a value of 0 implies that there is no presence of reference prices in the screen, whereas a value of 1 implies the presence of reference prices. This means that the Beta should be positive and significant in order to accept hypotheses 8a and 8b. When analyzing the three regression models of PU, it can be concluded that both conditions are not met; the Beta is negative and insignificant in all three models. The output of the three regression models of WtB are looking more

promising though. The Beta is positive and therefore in accordance with the hypothesized Beta in H8b. However, the effect of Reference Price in all three WtB models is found to be insignificant, with the highest p-value to be found in the full model (p = 0,317). This leads to the conclusion that both H8a and H8b are rejected.

4.3.4 Effect of Perceived Usability on Willingness to Buy

Next to the hypotheses that were developed for the eight independent screen variables, it was also hypothesized that the dependent variable perceived usability has a significant influence on the other dependent variable, willingness to buy. This led to the ninth hypothesis:

H9: Perceived usability positively influences willingness to buy

As the multiple regression models in section 4.2 focused on finding relationships between the independent and dependent variables, the effect of PU on WtB has not yet been analysed. In order to test for this relationship, a new multiple regression is created with willingness to buy as dependent variable and perceived as independent/predictor variable. The output of this can be found in the table below:

Table 11: Regression Model PU-WtB

Coefficients[a]

Model**		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	9,536	1,709		5,579	,000*
	Perceived Usability	,682	,026	,604	26,586	,000*

a. Dependent Variable: Willingness to Buy

1) * = statistically significant at 1%-level

2) ** = R Square of Model: 0,365, ANOVA: F: 706,806 -> p = 0,000*

Looking at this output, there are a few things that can clearly be noted. Firstly, the effect of perceived usability on willingness to buy is found to be highly significant at a 1%-level. The t value in this case is remarkably high. Secondly, the R^2 of the model is 0,365, which implies that perceived usability explains 36,5 percent of the variance of willingness to buy. This finding assumes that the model which includes PU as single predictor variable explains a larger part of the variance of WtB than the full regression model in part 4.2.2, which has an R^2 of 0,221. The last thing that can be concluded is that the effect of PU on WtB is of a positive nature, considering the positive Beta (0,682) of perceived usability. Therefore, the ninth and last hypothesis is accepted.

The conclusion of the testing of the various hypotheses is that several significant relationships have been found between variables, though due to the fact that the effects were often the opposite of the hypothesized effects, only three hypotheses are accepted. The other hypotheses are rejected. A summary of the hypotheses testing can be found in the table below.

Table 12: Hypotheses Testing

Hypothesis	Hypothesized Effect	Measured Effect	H0 accepted/rejected
1a: Screen Density -> PU	(-)	(+)*	Rejected
1b: Screen Density -> WtB	(-)	(+)**	Rejected
2a: Navigation Design -> PU	(-)	(-)	Rejected
2b: Navigation Design -> WtB	(-)	(+)	Rejected
3a: Symmetry -> PU	(+)	(-)**	Rejected
3b: Symmetry -> WtB	(+)	(-)	Rejected
4a: Color Design -> PU	(+)	(-)*	Rejected
4b: Color Design -> WtB	(+)	(-)*	Rejected
5a: Font Design -> PU	(+)	(+)	Rejected
5b: Font Design -> WtB	(+)	(-)	Rejected
6a: Professionalism -> PU	(-)	(+)	Rejected
6b: Professionalism -> WtB	(+)	(+)	Rejected
7a: Product Rating -> PU	(+)	(+)*	Accepted
7b: Product Rating -> WtB	(+)	(+)*	Accepted
8a: Reference Price -> PU	(+)	(-)	Rejected
8b: Reference Price -> WtB	(+)	(+)	Rejected
9: PU -> WtB	(+)	(+)*	Accepted

1) * = statistically significant at 1%-level
2) ** = marginally significant at 10%-level

5 Discussion

In the previous chapter it was concluded that only three hypotheses were supported. However, a total of six significant relationships and two marginally significant relationships were found. Although this still implies that in over half of the tested hypotheses there was no significant relationship, these result do provide some room for thought and discussion. Moreover, it is not unusual that not all the supposed relationships are proven to be significant as this is typically common for screening designs (Lawson, 2014). More interesting though is the fact that many of the relationships had opposite effects than the effects that were hypothesized. In this chapter these findings and the possible explanation of their occurrence is discussed. Furthermore, additional findings of this study are discussed.

5.1 Classical Aesthetics (CA)

For the three screen attributes that are associated with classical aesthetics, none of the seven hypothesis are supported. Moreover, five of the six hypothesized effects were actually found to be the opposite. Out of the three CA attributes, Navigation Design was clearly the least significant. Based on the results, it can be concluded that the navigation design of a screen doesn't significantly influence perceived usability and willingness to buy. However, it has to be noted that the differences in navigation design between simple and advanced design were possibly not denough to be clearly distinctive from one other. For future research, it could be useful to make this difference more extreme. Although the effects were not sufficient, one interesting thing could be noted in regard to the effect of navigation design on PU and WtB. The effect on willingness to buy is found to be negative, whereas the effect on perceived usability is found to be positive. This could be explained by the framing effect that was found in prior studies, in which the preference of users shifted when they were asked to perform a task (Hartmann et al., 2007; 2008; van Schaik & Ling, 2009), as 'buying' a product on a screen can be assumed to be more like a task than rating the usability of a screen. However, since the effects were not significant, the existence of such an effect should be examined in future studies.

In contrary to Navigation Design, the attribute Screen Density was found to have a highly significant effect on PU and WtB. The effect in both cases was positive though, whereas a negative effect was expected. This is an effect that hasn't been found in prior studies. The explanation for this effect could be that more products are associated with a larger product range. In the first part of the survey, respondents were asked to indicate the importance of aspects when shopping online. Although product range was rated as the lowest aspect (see Appendix A, figure 3E), results of the regression models showed that importance rating of product range has a significant effect on both PU and WtB. However, based on the data and output of this thesis, this effect can't be confirmed.

The last CA attribute that was tested in this study also showed effects that were exactly the opposite of what was hypothesized. Although the effect of Symmetry on PU was found to be marginally significant at a 10%-level, the effect on both PU and WtB was not significant at a 5%-level. However, based on the results, there is a suspicion that symmetry does have at least a small effect on perceived usability and willingness to buy. A possible explanation for this could be that the aesthetic judgement of symmetry on smartphone screens differ from that of larger screens, which would make sense because smartphone screens might be too small to symmetrically place products next to each other. Furthermore, the hypothesis that aesthetic judgements of male participants regarding symmetry are significantly different than those of female participants was rejected.

5.2 Expressive Aesthetics (EA)

Based on the results in chapter 4, the six hypotheses that are associated with expressive aesthetics were all rejected. However, it was found that Color Design has a significant effect on both PU and WtB, although the direction of the effect was not as expected. This difference could be explained by Hofstede's scale of individualism, on which the Netherlands score 80, which is considered to be very high. Cyr (2008) concluded that users from collectivist countries tend to have a preference for strong visuals, whereas users from individualistic countries (such as the Netherlands) tend to prefer a logical and structured page lay-out. However, there was an interesting additional finding for the Color Design attribute as the interaction variable MaleCD was found to be marginally significant at a 10%-level, with a positive Beta. This implies a suspicion that male respondents rate cool colors higher in terms of PU and WtB than their female counterparts. As this effect wasn't statistically confirmed at a 5%-level, it is very interesting to further examine this effect in future studies as this could provide valuable information for companies.

In contrary to Color Design, the other two EA attributes are found to be highly insignificant, implying that they don't have a notable influence on perceived usability or willingness to buy, although Professionalism seemed to have a small effect on PU. The direction of that effect was the opposite of the expected effect though. This could, again, be explained by the preference of individualistic cultures for logical and structured designs. In general however, it can be concluded that Font Design and Professionalism do not influence PU or WtB.

5.3 Product Specific Content (PSC)

The first of the two variables that is defined as product specific content, Product Ratings, is found to have a highly significant positive effect on both PU and WtB. These findings are in line with the two hypotheses for this variable. However, the question in this case is to what extend this effect is caused by the social influence effect as discussed by Burnkrant and Cousineau (1975) or by an increased aesthetic judgement due to the presence of product ratings. Based on the data and the output of this study, no such distinction can be revealed. It is interesting though to further explore this effect in future studies. The other variable that

belongs to the PSC category, Reference Price, was not found to have any significant effects on either PU or WtB.

5.4 Effect of Perceived Usability on Willingness to Buy

The last hypothesis that has been tested is the effect of PU on WtB, which was found to be highly significant according to the output in 4.3.4. Based on this, it can be concluded that the both dependent variables are highly correlated. This implies that a high perceived usability of a screen design leads to an increased willingness to buy certain products of that application. This is in accordance with the expected effect. However, one important thing has to be noted in this case. Some respondents whom I know personally gave me feedback on the survey, saying that it felt not rational to give a low score in the second question in the case that they gave a relative high rating in the first question. The rather extreme effect that was found in this thesis seems to confirm this problem. There is a clear suspicion that respondent used the first question of PU as anchor for their answer on the second question of WtB, which can lead to an anchoring effect. The anchoring effect can subsequently lead to biased data (Tversky & Kahneman, 1975). Unfortunately, the suspicion of the presence of an anchoring effect can't be confirmed, nor can it be measured what the effect of PU on WtB would be without the presence of an anchoring effect. Therefore, it is assumed that perceived usability does have a significant influence on willingness to buy, but with the remark that the measured effect is possibly larger than in reality due to an anchoring effect.

5.5 Additional Findings

Next to the eight screen attributes, control variables were added to find any possible confounding effects and to check for the correctness of some assumptions. Several of these control variables are found to have a significant influence, although these influences differ for both dependent variables. One of those control variables is Frequency of buying Online. This variable has a positive significant effect on PU, whereas it has a negative, highly insignificant effect on WtB. This finding implies that the more frequent respondents buy products online, the higher they perceive the usability of the different screen designs. A possible explanation for this is that people who frequently buy online are more familiar with online user interfaces and are better in finding their way within those interfaces. This would lead these respondents to rate designs, in general, higher on usability.

Another interesting finding was the effect of gender. In the case of perceived usability as dependent variable, the effect of gender was found to be minimal, implying that screens are on average rated as equally usable by female and male. However, when looking at willingness to buy as dependent variable, this is considerably different. Based on the extended regression model of WtB, it can be concluded that male respondents are considerably less willing to buy a food processor. However, the full model with interaction variables gives a complete new view on this. It can be derived from this model that this difference is for a major part caused by the fact that male respondents who already own a

food processor give a considerably lower rating of WtB. Next to that, gender is also included in the interaction variable MaleCD. When correcting for these interaction variables, the control variable gender is not significant anymore, implying that on average there is no significant difference between male and female in terms of willingness to buy.

Furthermore, it is interesting to see if it makes any difference whether a respondent is already in possession of a food processor, whether he or she is currently interested in buying a food processor and to check for combinations of those dummy variables. Purely looking at the effect of these dummy variables without interaction effects (i.e. the effects in the extended regression models), a clear difference can be noted between the model of PU and the model of WtB. In the first model, neither Food Processor Ownership nor Interest in buying a Food Processor has a significant effect, whereas the effect of both variables on WtB is highly and positively significant. This implies that respondents who already own a food processor rate the designs higher on willingness to buy. A possible explanation for this could be that respondents who are currently in possession of a food processor generally have more interest in the product than those who are not in possession of a food processor, which translates itself in higher ratings of WtB. Not surprisingly, the respondents who are currently interested in buying a food processor are more likely to give higher ratings of WtB.

However, when the interaction variable OwnFP_BuyFP is included, things suddenly got much more interesting. The inclusion of this interaction variable increased the effect of the two separate dummy variables. Moreover though, the effect of the interaction variable itself is highly, negatively significant. This implies that respondents who already own a food processor, but are still interested in buying a food processor, are giving considerably lower ratings of willingness to buy for the different screen designs. Next to that, when adding this interaction variable in the PU model, the same effect can be noted. The separate effects of Food Processor Ownership and Interest in buying a Food Processor on PU also got highly significant after the interaction variable was added. This means that respondents who already own a food processor, but are still interested in buying a food processor, give overall lower ratings for both dependent variables. This finding can be explained by the fact that these type of respondents are highly critical when it comes to buying a new food processor. They already own a food processor, but they might probably be looking for a better one, which makes them evaluate food processors more critical.

6. Conclusions

This chapter focuses on answering the research question that was stated in chapter 1, based on the conclusions of the various hypotheses. Next to this general discussion, academic contributions and managerial implications are discussed. Last but not least, limitations of this research and directions for future research are discussed.

6.1 General Conclusion

In chapter 1, it has been argued that understandings of the effect of design aesthetics are limited, especially the understanding of separate screen attributes on perceived usability and willingness to buy. This thesis attempted to provide more insight in these specific effects, based on the following central research question:

Can design aesthetics influence the perceived usability of mobile websites and applications and can it subsequently increase the customer's willingness to buy?

In order to answer this research question, 18 hypotheses were developed. Although only three of these hypotheses were supported, a total of six significant relationships and two marginally significant relationships were found, which is a reasonable outcome for a study with a screening design (Lawson, 2014). One of the most interesting findings of this study is that none of the hypothesized effects that are associated with either classical or expressive aesthetics was supported. Does this imply that design aesthetics do not influence the perceived usability of mobile websites and applications and the willingness to buy from those sites? If no significant effects were found at all, this would have indeed been the conclusion.

However, although none of the aesthetic related hypotheses were supported, the results showed significant effects. The fact that the direction of these effects is opposite to what was hypothesized doesn't mean that there is no influence of these attributes on PU and WtB. The two screen attributes that were found to be most influential are the screen density (CA attribute) of an application and the color design (EA attribute) of an application. Furthermore, the symmetry of the placements of products on a screen is also thought to have an impact on the perceived usability of an application. Based on these findings, it can indeed be concluded that design aesthetics have an influence on the perceived usability of mobile websites and applications and that it affects the willingness to buy products from those retailers. Furthermore, the presence of product rating has been found to have a significant influence on both PU and WtB, although it is difficult to determine whether this is caused by the social influence effect of product ratings or an increased aesthetical judgement by showing these ratings. The last conclusion that can be drawn from this study is that perceived usability and willingness to buy are highly correlated, implying that designs with a high perceived usability lead to an increased willingness to buy.

A main objective of this thesis was to apply the findings of the study to find the optimal aesthetic design of mobile applications, based on the eight examined screen attributes. In order to find this optimal screen design, the direction of the effects that were measured in the full regression models of PU and WtB are taken as indicators for the preference of the respondents. The preferences of the respondents for both an optimal usability and willingness to buy can be found in table 2 of appendix A. It can be noted that the optimal screen in terms of PU differs from the optimal screen design for WtB. The attributes for which these differences occur were actually the three attributes that were found to be least significant though, so it is rather the question if these difference does truly exist. As the ultimate goal of companies is to get users to buy their products, the optimal screen that is chosen is the one for willingness to buy. In figure 4 (left) of appendix A, this screen has been designed in order to give a more graphical idea of how the optimal screen looks like. Next to this screen another screen is designed which is, based on the results of this study, assumed to be the least optimally designed screen (figure 4 right). Implications of these most and least optimal designs are discussed in paragraph 3 of this chapter.

6.2 Academic Contribution

From an academic point of view, this study provides contributions in several ways. Firstly, the extensive literature review of this study combines knowledge of usability, (mobile) user interface design and design aesthetics in such a way that the topic of design aesthetics can be understood through the underlying usability issues. Although there are numerous studies that have researched usability issues in m-commerce (i.a. Tarasewich, 2003a; Shankar & Balasubramanian, 2009; Shankar et al., 2010), none of these studies have applied these issues to user interface design or design aesthetics. By doing so in this study, one gets a better understanding of the behavior of consumers in the mobile environment. This way of combining prior studies, combined with the results of this study, is a first step in developing a solid framework for mobile interface designers. Although design aesthetics are just a part of the mobile interface, this study does contribute at least for a part in developing such a framework.

Secondly, an important objective of this research from an academic point of view was to separately examine the effect of various screen attributes on perceived usability and willingness to buy. Although some prior studies have made distinctions between different types of aesthetics (Cyr, 2008; van Schaik & Ling, 2009) and some studies have examined the effect of just one of the attributes (Bauerly & Liu, 2006; Tuch et al., 2010), there is no prior research in which the effect of separate attributes of design aesthetics is conjointly examined. Results of this study provide evidence that there are indeed differences between the effects of various attributes of design aesthetics, with some having no significant influence and others having a highly significant influence.

The last academic contribution of this study is the fact that it found a couple of significant effects that were exactly the opposite of the hypothesized effects. Although the general conclusion is in line with the expectations that are based on prior research, the separate hypotheses considerably differ from the actual findings. Especially the findings on the effects of screen density and color design on PU and WtB are interesting, as these findings are strongly in contrast with previous findings on this matter. Although existing literature provides a possible explanation for the different effect of color design, the surprising results of this study shed new light on these variables which has not been found before. It is definitely something that has to be further investigated in future academic research.

6.3 Managerial Implications

This thesis started with pointing out that the conversion rates of Asian countries in the smartphone industry are considerably higher than those of retailers in western countries. Although it was suggested that this could have partly been the consequence of differences in cultural dimensions which affect the adoption rates of new technology, evidence was found that the main problem in this case are the (usability) issues in the m-commerce environment. As a result of this finding, the possibility was suggested that the difference in conversion rates could be the consequence of difference in mobile interface designs between Asian and western countries. In chapter 3, it was concluded that there are indeed differences in design between the two cultures. This fact, in combination with extensive literature review, led to the identification of eight variables, which were tested in an experiment. The main objective of this was to find the optimal combination of these eight screen attributes. However, a second objective from a managerial point of view was to investigate if this optimal mobile interface design can explain the differences in conversion rates.

In the first section of this chapter, an optimized combination of screen attributes has been developed, based on the results of the experiment. Knowledge of how to optimally design the mobile interface in terms of design aesthetic is very valuable for companies. This study showed that design aesthetics can lead to a higher perceived usability of a site or application and can also increase user's willingness to buy. This could on its turn increase the conversion rates and ultimately the profit of retailer's mobile commerce activities. Furthermore, this study provided evidence that effects of certain screen attributes are not consistent in all scenarios. For example, in general a preference was found for a neutral color design. However, when this effect was further investigated, it was found that male and female tend to have different preferences regarding color design. From a managerial point of view, this implies that segmentation of customers and knowing the preferences of customers is vital when designing a mobile website or application. Likewise it is also important to search for any differences in terms of feelings towards certain products, although this is something which is not related to design aesthetics. Nevertheless, results of this study showed that

these kind of factors can also influence the perceived usability and willingness to buy. Therefore, it is important to also segment customers based on this kind of information.

To conclude the implications for managers, it is interesting to apply the optimal combination of attributes to the question that was raised in the beginning of this thesis. Does Asian retailers indeed have a better understanding of the dynamics of mobile commerce, which explains the higher conversion rates in these countries? This question can't be answered merely by the results of this study, since this study focused purely on design aesthetics. However, if this study indeed shows evidence that Asian screen designs are perceived as more user friendly and that they increase user's willingness to buy, that would be a first confirmation of this suspicion. In order to test this suspicion, the screenshots of Rakuten and Amazon are compared with the most optimal and least optimal screen designs (figure 2 and 4 of Appendix A). Based on the abovementioned suspicion, one would expect the design of the Rakuten application to have high resemblance with the most optimal screen design.

When comparing the four screens, the findings are very interesting. The most optimal screen design is almost identical to the Amazon screen design, whereas the Rakuten screen design has high resemblance with the least optimal screen design. This finding can be interpreted in two ways. Firstly, this conclusion is in accordance with the aforementioned findings of Cyr (2008), who found that users from individualistic countries tend to prefer a logical and structured design above a strong visual design. Secondly though, this finding is completely in contrast with the suspicion that the difference between the conversion rates could be caused by difference in screen designs. Therefore, it can be concluded that aesthetical differences in screen designs between eastern and western cultures are not the cause of the difference in conversion rates between the two cultures. However, it has to be noted that this study only focused on the effect of design aesthetics in the western culture. It would be interesting to conduct the exact same study in a collectivist culture in order to compare the results of both studies. Clearly, further research is needed to reveal the specific dynamics of the mobile commerce environment. Directions for future research are discussed in the next section.

6.4 Limitations and Directions for Future Research

Throughout this thesis, several limitations and directions for future research have already been briefly mentioned. In this section these and other limitations are discussed more extensively, along with suggestions for directions for future research. For the sake of simplicity, limitations are separated into limitations of internal validity and limitations of external validity.

6.4.1 Internal Validity

In order to find the effect of design aesthetics on perceived usability and willingness to buy, an experiment was conducted with an online survey as method to collect the data. Although one typically has more control over the internal validity in these type of experiments, a couple of limitations in terms of internal validity were encountered during this thesis. A first limitation could lie in the manipulations of the eight different screen attributes. The manipulations were mostly based on the differences in screen design between the applications of Rakuten and Amazon, which would increase the external validity of research. However, some respondents gave me feedback that they were not able to identify clear differences between the screen. The attributes that were mostly identified by these respondents as being different among the screen designs were image size (which in this case is represented by screen density) and color. It is therefore not surprising that these two attributes are the attributes that were found to have a significant effect on PU and WtB. One could say that the manipulations used in the experiment resemble real screen designs and are therefore well chosen. However, in these kind of experiments it could also be interesting to make the manipulations more extreme. In that case the influence of other, less prominent attributes, can be measured more precisely. Applying these kind of extreme manipulations could be interesting for future research.

More limitations could lie in the measurement of the dependent variables. To measure the dependent variables, a single-item measure scale was used. This measure scale was chosen because of the fact that the respondents were thought to have a limited amount of time. In those cases, single-item measure scales are a good alternative for multi-item measure scales (Waltz et al., 1991). Although prior research has found that single-item measure scales can be sufficient in order to capture certain responses (Bergkvist & Rossiter, 2007; Wanous et al., 1997), other studies dispute the use of single-item scales as they are unable to measure different dimensions of a variable (Youngblut & Casper, 1993). For future research it could therefore be interesting to apply multi-item measure scales like the one used in the study of Cyr et al. (2006) in order to capture effects on different dimensions of especially perceived usability. A major requirement to conduct such a study though is that the respondents are not busy. Since the amount of time to complete a study in that case would probably be considerably higher, it might be better to not distribute the surveys online. Another limitation that is linked to the used measurement scale is the possible presence of an anchoring effect. The suspicion of the presence of this effect has already been discussed in section 5.4, although the actual presence of the effect can't be confirmed. Next to being able to measure different dimensions of dependent variables, applying multi-item measurement scales in future studies can also considerably decrease anchoring effects. This is a lesson that has to be taken into account when designing similar studies in the future.

6.4.2 External Validity

Next to limitations of internal validity, it can be argued that this thesis has several limitations in terms of external validity as well. A major limitation of external validity in this study is the fact that, since the data were collected with online surveys, only the subjective responses of the respondents could be measured. Especially for the dependent variable perceived usability, this could be problematic for the external validity of the study. Studies of Hartmann, Sutcliffe and de Angeli (2007; 2008) has showed that there are significant differences between subjective and objective usability in studies focused on design aesthetics. Moreover, Duh, Tan and Chen (2006) found that the setting in which usability testing is done can yield significantly different results. They suggest that this is caused by the fact that some usability problems are only related to the device being used in the field. These problems can't be found in conventional laboratory usability tests. This thesis can be defined as a laboratory test, which implies that more interactive, field application of this study could produce significantly different results. Conducting such an experiment can be very interesting for future research.

Another source of limitations of external validity could be the sample of the study. The sample of this study consisted of Dutch students. Although it is not always necessary to have a sample that consists of a fair representation of the entire population, there are two limitations with this sample. The first and most important limitation is the lack of cultural difference in the sample. In section 6.3 it was already concluded that the results of this study possibly depend on the cultural dimension of the respondents. According to Hofstede's scale of individualism, the Netherlands is a highly individualistic country. As concluded before in this study, this could explain the fact that respondents prefer a logical and structured screen design without strong visuals. Although conclusions can be drawn for western countries, it is rather the question whether the same experiment in a strongly collectivist country yields the same results. Therefore it is interesting to conduct a study that accounts for differences in cultural dimensions, as this could significantly change the implication of this study.

The second limitation that stems from this sample is the fact that the sample is focused on students. Dutch students are typically highly educated people in the age group 18-25, with a relatively low income. This is quite a specific part of the population and therefore the results of this study can't be generalized for the entire population of western countries. This study did not have the time nor the resources to collect data from a sample that was representative for the entire population. In order to get a complete understanding of the dynamics of mobile commerce and the effect of design aesthetics on perceived usability and willingness to buy, it would be necessary to conduct an experiment that includes such a diverse sample. However, despite of the limitations this study provides valuable information for both scientific and managerial purposes. It should serve as a contribution towards a full understanding of mobile commerce.

7 References

Acquisti, A., & Spiekermann, S. (2011). Do interruptions pay off? Effects of interruptive ads on consumers' willingness to pay. *Journal of Interactive Marketing*, *25*(4), pp. 226-240.

Adipat, B., & Zhang, D. (2005). Interface design for mobile applications.*AMCIS 2005 Proceedings*, 494.

Akhtar, O. (2013, March 22). Rakuten: The biggest e-commerce site you haven't heard of. Retrieved from http://fortune.com/2013/03/22/rakuten-the-biggest-e-commerce-site-you-havent-heard-of/

Al-Alawi, A. I. (2006). Wi-Fi technology: Future market challenges and opportunities. *Journal of computer Science*, *2*(1), pp. 13-18.

Anderson, E. W. (1998). Customer satisfaction and word of mouth. *Journal of service research*, *1*(1), pp. 5-17.

Anderson, M. (2015a, April 29). For vast majority of seniors who own one, a smartphone equals 'freedom'. Retrieved from http://www.pewresearch.org/facttank/2015/04/29/seniors-smartphones/.

Anderson, M. (2015b, October 29). Technology Device Ownership: 2015. Retrieved from http://www.pewinternet.org/2015/10/29/technology-device-ownership-2015/

AT&T. (2015, May). Smartphone Use Behind the Wheel Survey. Retrieved from http://about.att.com/content/dam/snrdocs/2015%20It%20Can%20Wait%20Report_Smartphone%20Use%20Behind%20the%20Wheel%20.pdf.

Babin, B. J., Hardesty, D. M., & Suter, T. A. (2003). Color and shopping intentions: The intervening effect of price fairness and perceived affect. *Journal of Business Research*, *56*(7), pp. 541-551.

Bachiochi, D., Berstene, M., Chouinard, E., Conlan, N., Danchak, M., Furey, T., ... & Way, D. (1997). Usability studies and designing navigational aids for the World Wide Web. *Computer Networks and ISDN systems*, *29*(8), pp. 1489-1496.

Balasubramanian, S., Peterson, S. A. & Jarvenpaa, S. L. (2002). Exploring the Implications of M-Commerce for Markets and Marketing. *Journal of Academy of Marketing Science, 30* (4), pp. 348–61.

Bannan, K. J. (2000). The promise and perils of wap. *Scientific American*,*283*(4), pp. 46-49.

Barnes, S. J. (2002). Provision of services via the wireless application protocol: a strategic perspective. *Electronic Markets*, *12*(1), pp. 14-21.

Barredo, A. (2014, May 28). A comprehensive look at smartphone screen size statistics and trends. Retrieved from https://medium.com/@somospostpc/a-comprehensive-look-at-smartphone-screen-size-statistics-and-trends-e61d77001ebe#.t8z4vfico.

Baty, J.B. II & Lee, R.M. (1995). Intershop: Enhancing the vendor/customer dialectic in electronic shopping. *J. Manage. Info. Syst. 11*, 4, 9–31.

Bauerly, M., & Liu, Y. (2006). Computational modeling and experimental investigation of effects of compositional elements on interface and design aesthetics. *International Journal of Human-Computer Studies*, *64*(8), pp. 670-682.

Beach, A., Gartrell, M., & Han, R. (2009, August). Solutions to security and privacy issues in mobile social networking. In *Computational Science and Engineering, 2009. CSE'09. International Conference on* (Vol. 4, pp. 1036-1042). IEEE.

Bellizzi, J. A., & Hite, R. E. (1992). Environmental color, consumer feelings, and purchase likelihood. *Psychology & marketing*, *9*(5), pp. 347-363.

Bergkvist, L., & Rossiter, J. R. (2007). The predictive validity of multiple-item versus single-item measures of the same constructs. *Journal of marketing research*, *44*(2), pp. 175-184.

Bigné, E., Ruiz, C., & Sanz, S. (2007). Key drivers of mobile commerce adoption. An exploratory study of Spanish mobile users. *Journal of theoretical and applied electronic commerce Research*, *2*(2), pp. 48-60.

Biswas, A., & Blair, E. A. (1991). Contextual effects of reference prices in retail advertisements. *The Journal of Marketing*, 1-12.

Blattberg, R. C., & Neslin, S. A. (1990). *Sales promotion: Concepts, methods, and strategies*. Ney Jersey: Prentice Hall.

Bonnington, C. (2013, November 4). Smartphone Screen Sizes Keep On Growing — But Not for Much Longer. Retrieved from http://www.wired.com/2013/04/why-big-smartphone-screens/.

Boshoff, D. (2014, April 4). Smartphone adoption in Africa: The rising smartphone market. Retrieved from http://www.amgoo.com/blog/smartphone-adoption-in-africa-the-rising-smartphone-market

Brown, K. V. (2015, March 9). Researchers say people are choosing public transit so they can stare at their smartphones. Retrieved from http://fusion.net/story/192808/public-transit-is-the-greatest-smartphone-fix-time/.

Bruner, G. C., & Kumar, A. (2005). Explaining consumer acceptance of handheld Internet devices. *Journal of business research*, *58*(5), pp. 553-558.

Buellingen, F., & Woerter, M. (2004). Development perspectives, firm strategies and applications in mobile commerce. *Journal of Business Research*, *57*(12), pp. 1402-1408.

Bulmer, M. G. (2012). *Principles of statistics*. Courier Corporation.

Burnkrant, R. E., & Cousineau, A. (1975). Informational and normative social influence in buyer behavior. *Journal of Consumer research*, *2*(3), pp. 206-215.

Cha, M. (2015, March 13). What Is 5G, and What Does It Mean for Consumers?. Retrieved from http://www.recode.net/2015/3/13/11560156/what-is-5g-and-what-does-it-mean-for-consumers

Chan, S. S., Fang, X., Brzezinski, J. R., Zhou, Y., Xu, S., & Lam, J. (2002). Usability for Mobile Commerce Across Multiple Form Factors. *J. Electron. Commerce Res.*, *3*(3), pp. 187-199.

Chaudry, B. M., Connelly, K. H., Siek, K. A., & Welch, J. L. (2012, January). Mobile interface design for low-literacy populations. In *Proceedings of the 2nd ACM SIGHIT international health informatics symposium* (pp. 91-100). ACM.

Chen, Y. (2016, January 25). Amazon, Alibaba and Rakuten: who is winning the global ecommerce game?. Retrieved from https://www.clickz.com/amazon-alibaba-and-rakuten-who-is-winning-the-global-ecommerce-game/91882/

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

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

Childers, T. L., Carr, C. L., Peck, J., & Carson, S. (2002). Hedonic and utilitarian motivations for online retail shopping behavior. *Journal of retailing*,*77*(4), pp. 511-535.

Choi, J. H., & Lee, H. J. (2012). Facets of simplicity for the smartphone interface: A structural model. *International Journal of Human-Computer Studies*, *70*(2), pp. 129-142.

Cleff, B. E. (2007). Privacy issues in mobile advertising. *International Review of Law Computers and Technology*, *21*(3), 225-236.

Coursaris, C., & Hassanein, K. (2002). Understanding m-commerce: a consumer-centric model. *Quarterly journal of electronic commerce*, *3*, pp. 247-272.

Criteo. (2015). State of Mobile Commerce. Growing like a weed. *Q1 2015*

Crowley, A. E. (1993). The two-dimensional impact of color on shopping. *Marketing letters*, *4*(1), pp. 59-69.

Curtis, S. (2013, October 10). Quarter of mobile users fell victim to cybercrime last year. Retrieved from http://www.telegraph.co.uk/technology/internet-security/10369081/Quarter-of-mobile-users-fell-victim-to-cybercrime-last-year.html

Cyr, D. (2008). Modeling web site design across cultures: relationships to trust, satisfaction, and e-loyalty. *Journal of Management Information Systems*, *24*(4), pp. 47-72.

Cyr, D., Head, M., & Ivanov, A. (2006). Design aesthetics leading to m-loyalty in mobile commerce. *Information & Management*, *43*(8), pp. 950-963.

Dai, H., & Palvia, P. C. (2009). Mobile commerce adoption in China and the United States: a cross-cultural study. *ACM SIGMIS Database*, *40*(4), pp. 43-61.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS quarterly*, pp. 319-340.

Dellarocas, C., Zhang, X. M., & Awad, N. F. (2007). Exploring the value of online product reviews in forecasting sales: The case of motion pictures.*Journal of Interactive marketing*, *21*(4), pp. 23-45.

Detenber, B. H., & Reeves, B. (1996). A bio-informational theory of emotion: Motion and image size effects on viewers. *Journal of Communication*, *46*(3), pp. 66-84.

Deutsch, M. & Gerard, H. B. (1955). A study of normative and informational social influences upon individual judgment. *Journal of Abnormal and Social Psychology*, *51*, pp. 629-636.

Dholakia, R. R., & Dholakia, N. (2004). Mobility and markets: emerging outlines of m-commerce. *Journal of Business research*, *57*(12), pp. 1391-1396.

Digital Trends. (2016, June 7). Who needs an umbrella? These waterproof Android phones can handle the rain. Retrieved from http://www.digitaltrends.com/mobile/best-waterproof-phones/2/#:kZHpop4f0nbutA

Duh, H. B. L., Tan, G. C., & Chen, V. H. H. (2006, September). Usability evaluation for mobile device: a comparison of laboratory and field tests. In*Proceedings of the 8th conference on Human-computer interaction with mobile devices and services* (pp. 181-186). ACM.

Dunlop, M., & Brewster, S. (2002). The challenge of mobile devices for human computer interaction. *Personal and ubiquitous computing*, *6*(4), pp. 235-236.

EcommerceNews. (2015, June 18). Top 500 European e-retailers generated sales of €124 billion. Retrieved from http://ecommercenews.eu/top-500-european-e-retailers-generated-sales-of-e124-billion/.

Elliott, A. C., & Woodward, W. A. (2007). *Statistical analysis quick reference guidebook: With SPSS examples*. Sage.

eMarketer (2015, January 8). Tablet Users to Surpass 1 Billion Worldwide in 2015. Retrieved from http://www.emarketer.com/Article/Tablet-Users-Surpass-1-Billion-Worldwide-2015/1011806

Eroglu, S. A., Machleit, K. A., & Davis, L. M. (2003). Empirical testing of a model of online store atmospherics and shopper responses. *Psychology & Marketing*, *20*(2), pp. 139-150.

Fang, X., & Holsapple, C. W. (2007). An empirical study of web site navigation structures' impacts on web site usability. *Decision Support Systems*, *43*(2), pp. 476-491.
Flavián, C., Guinalíu, M., & Gurrea, R. (2006). The role played by perceived usability, satisfaction and consumer trust on website loyalty. *Information & Management*, *43*(1), 1-14.

Gartner Research (2011). Forecast: Mobile Advertising, Worldwide, 2008-2015.

Ghosh, A. K., & Swaminatha, T. M. (2001). Software security and privacy risks in mobile e-commerce. *Communications of the ACM*, *44*(2), 51-57.

Goldman, D. (2015, May 19). 5G will cost you a bundle. Retrieved from http://money.cnn.com/2015/05/18/technology/5g-cost-wireless-data/

Gong, J., & Tarasewich, P. (2004, November). Guidelines for handheld mobile device interface design. In *Proceedings of DSI 2004 Annual Meeting* (pp. 3751-3756).

Gorn, G. J., Chattopadhyay, A., Sengupta, J., & Tripathi, S. (2004). Waiting for the web: how screen color affects time perception. *Journal of marketing research*, *41*(2), pp. 215-225.

Grewal, D., Krishnan, R., Baker, J., & Borin, N. (1998a). The effect of store name, brand name and price discounts on consumers' evaluations and purchase intentions. *Journal of retailing*, *74*(3), 331-352.

Grewal, D., Monroe, K. B., & Krishnan, R. (1998b). The effects of price-comparison advertising on buyers' perceptions of acquisition value, transaction value, and behavioral intentions. *The Journal of Marketing*, 46-59.
GSMA Intelligence. (2015, July). Rural coverage: strategies for sustainability: country case studies.

H. Bauer, H., Reichardt, T., Exler, S., & Tranka, E. (2007). Utility-based design of mobile ticketing applications–a conjoint-analytical approach. *International Journal of Mobile Communications*, *5*(4), pp. 457-473.

Hall, R. H., & Hanna, P. (2004). The impact of web page text-background colour combinations on readability, retention, aesthetics and behavioural intention. *Behaviour & information technology*, *23*(3), pp. 183-195.

Hartmann, J., Sutcliffe, A., & De Angeli, A. (2007). Investigating attractiveness in web user interfaces. *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 387-396. ACM.

Hartmann, J., Sutcliffe, A., & De Angeli, A. (2008). Towards a theory of user judgment of aesthetics and user interface quality. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *15*(4), 15.

Helson, H. (1964). *Adaptation-level theory*. Oxford, England: Harper & Row Adaptation-level theory.
Herzberg, A. (2003). Payments and banking with mobile personal devices.*Communications of the ACM*, *46*(5), pp. 53-58.

Hill, A., & Scharff, L. V. (1997). Readability of websites with various foreground/background color combinations, font types and word styles. In*Proceedings of 11th National Conference in Undergraduate Research* (Vol. 2, pp. 742-746).

Hofstede, G. (1984). Cultural Dimensions In Management And Planning. *Asia Pacific Journal of Management, January,* pp. 81-99

Hofstede, G. (1997). *Cultures and Organizations: Software of the Mind*. New York: McGraw-Hill.

Ivan, I., Milodin, D., & Zamfiroiu, A. (2013). Security of M-Commerce transactions. *Theoretical And Applied Economics*, *20*(7), 59-76.

Juul, N. C., & Jørgensen, N. (2003). The security hole in WAP: an analysis of the network and business rationales underlying a failure. *International Journal of Electronic Commerce*, *7*(4), pp. 73-92.

Kane, S. K., Bigham, J. P., & Wobbrock, J. O. (2008, October). Slide rule: making mobile touch screens accessible to blind people using multi-touch interaction techniques. In *Proceedings of the 10th international ACM SIGACCESS conference on Computers and accessibility* (pp. 73-80). ACM.

Keith, T. Z. (2014). *Multiple regression and beyond: An introduction to multiple regression and structural equation modeling*. Routledge.

Kelman, H. (1958). Compliance, identification, and internalization: Three processes of attitude change. *Journal of Conflict Resolution*, *1*, pp. 51-60.

Khan, M. A. R. (2014). Investigation for the Wireless Communication and intended about the 5th Generation Technology. *International Journal of Innovation and Applied Studies*, *7*(1), 401.

Kim, E. B., Eom, S. B. (2002). Designing effective cyber store user interface. *Industrial Management & Data Systems, Vol. 102,* Iss 5, pp. 241 - 251.

Kim, H., Kim, J., Lee, Y., Chae, M., & Choi, Y. (2002, January). An empirical study of the use contexts and usability problems in mobile Internet. In *System Sciences, 2002. HICSS. Proceedings of the 35th Annual Hawaii International Conference on* (pp. 1767-1776). IEEE.

Kim, J., & Moon, J. Y. (1998). Designing towards emotional usability in customer interfaces—trustworthiness of cyber-banking system interfaces. *Interacting with computers*, *10*(1), pp. 1-29.

Kobayashi, M., Hiyama, A., Miura, T., Asakawa, C., Hirose, M., & Ifukube, T. (2011, September). Elderly user evaluation of mobile touchscreen interactions. In *IFIP Conference on Human-Computer Interaction* (pp. 83-99). Springer Berlin Heidelberg.

Kondratova, I., & Goldfarb, I. (2007, July). Color your website: Use of colors on the web. In *International Conference on Usability and Internationalization* (pp. 123-132). Springer Berlin Heidelberg.

Koo, D. M., & Ju, S. H. (2010). The interactional effects of atmospherics and perceptual curiosity on emotions and online shopping intention. *Computers in Human Behavior*, *26*(3), pp. 377-388.

Kopalle, P. K., & Lindsey-Mullikin, J. (2003). The impact of external reference price on consumer price expectations. *Journal of Retailing*, *79*(4), 225-236.

Koufaris, M. (2002). Applying the technology acceptance model and flow theory to online consumer behavior. *Information systems research*, *13*(2), pp. 205-223.

Krajnc, E., Knoll, M., Feiner, J., & Traar, M. (2011, November). A touch sensitive user interface approach on smartphones for visually impaired and blind persons. In *Symposium of the Austrian HCI and Usability Engineering Group* (pp. 585-594). Springer Berlin Heidelberg.

Kumar, V., Chattaraman, V., Neghina, C., Skiera, B., Aksoy, L., Buoye, A., & Henseler, J. (2013). Data-driven services marketing in a connected world.*Journal of Service Management*, *24*(3), 330-352.

Kurosu, M., & Kashimura, K. (1995, May). Apparent usability vs. inherent usability: experimental analysis on the determinants of the apparent usability. In *Conference companion on Human factors in computing systems* (pp. 292-293). ACM.

Latha, D. H., Reddy, D. R. K., Sudha, K., Mubeen, A., & Savita, T. S. (2014). A Study on 5th Generation Mobile Technology-Future Network Service. *IJCSIT) International Journal of Computer Science and Information Technologies*, *5*(6), 8309-8313.

Lattin, J. M., & Bucklin, R. E. (1989). Reference effects of price and promotion on brand choice behavior. *Journal of Marketing research*, 299-310.

Lavie, T., & Tractinsky, N. (2004). Assessing dimensions of perceived visual aesthetics of web sites. *International journal of human-computer studies*, *60*(3), pp. 269-298.

Lawson, J. (2003). One-step screening and process optimization experiments. *The American Statistician*, *57*(1), pp. 15-20.

Lawson, J. (2014). *Design and Analysis of Experiments with R*. CRC Press

Lee, Y. E., & Benbasat, I. (2003). Interface design for mobile commerce. *Communications of the ACM*, *46*(12), pp. 48-52.

Lee, Y. E., & Benbasat, I. (2004). A framework for the study of customer interface design for mobile commerce. *International Journal of Electronic Commerce*, *8*(3), pp. 79-102.

Lehr, W., & McKnight, L. W. (2003). Wireless internet access: 3G vs. WiFi?.*Telecommunications Policy*, *27*(5), pp. 351-370.

Li, Y. M., & Yeh, Y. S. (2010). Increasing trust in mobile commerce through design aesthetics. *Computers in Human Behavior*, *26*(4), pp. 673-684.

Lin, H. H., & Wang, Y. S. (2006). An examination of the determinants of customer loyalty in mobile commerce contexts. *Information & management*,*43*(3), pp. 271-282.

Lohse, G. L., & Spiller, P. (1998a). Quantifying the effect of user interface design features on cyberstore traffic and sales. *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 211-218. ACM Press/Addison-Wesley Publishing Co.

Lohse, G. L., & Spiller, P. (1998b). Electronic shopping. *Communications of the ACM*, *41*(7), 81-87.

MacInnes, I., Moneta, J., Caraballo, J., & Sarni, D. (2002). Business models for mobile content: The case of m-games. *Electronic Markets*, *12*(4), pp. 218-227.

Magura, B. (2003). What hooks m-commerce customers. *MIT Sloan Management Review*, *44*(3), 9.

Marcus, A., & Gould, E. W. (2000). Crosscurrents: cultural dimensions and global Web user-interface design. *interactions*, *7*(4), 32-46.

McCarthy, N. (2014, June 25). Reasons for Shopping on a Computer vs. Mobile Device. Retrieved from https://www.statista.com/chart/2392/reasons-for-shopping-on-a-computer-vs-mobile-device/

McCoy, S., Everard, A., & Loiacono, E. T. (2009). Online ads in familiar and unfamiliar sites: Effects on perceived website quality and intention to reuse. *Information Systems Journal*, *19*(4), pp. 437-458.

McCoy, S., Everard, A., Polak, P., & Galletta, D. F. (2007). The effects of online advertising. *Communications of the ACM*, *50*(3), pp. 84-88.

Moe, W. W., & Trusov, M. (2011). The value of social dynamics in online product ratings forums. *Journal of Marketing Research*, *48*(3), pp. 444-456.

Moshagen, M., & Thielsch, M. T. (2010). Facets of visual aesthetics. *International Journal of Human-Computer Studies*, *68*(10), pp. 689-709.

Mousa, A. M. (2012). Prospective of fifth generation mobile communications.*International Journal of Next-Generation Networks (IJNGN)*, *4*(3), 1-30.

Nanda, P., Bos, J., Kramer, K. L., Hay, C., & Ignacz, J. (2008). Effect of smartphone aesthetic design on users' emotional reaction: An empirical study.*The TQM Journal*, *20*(4), pp. 348-355.

Niranjanamurthy, M., Kavyashree, N., Jagannath, S., & Chahar, D. (2013). Analysis of e-commerce and m-commerce: advantages, limitations and security issues. *International Journal of Advanced Research in Computer and Communication Engineering*, *2*(6). pp. 2360-2370.

Norman, D. (2002). Emotion & design: attractive things work better.*interactions*, *9*(4), pp. 36-42.

Nunes, L., & Recarte, M. A. (2002). Cognitive demands of hands-free-phone conversation while driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, *5*(2), pp. 133-144.

O'Boyle, B. (2015, October 12). What is Siri? Apple's personal voice assistant explained. Retrieved from http://www.pocket-lint.com/news/112346-what-is-siri-apple-s-personal-voice-assistant-explained.

Osborne, J., & Waters, E. (2002). Four assumptions of multiple regression that researchers should always test. *Practical assessment, research & evaluation*,*8*(2), pp. 1-9.

Öztuna, D., Elhan, A. H., & Tüccar, E. (2006). Investigation of four different normality tests in terms of type 1 error rate and power under different distributions. *Turkish Journal of Medical Sciences*, *36*(3), 171-176.

Paradiso, J. A., & Starner, T. (2005). Energy scavenging for mobile and wireless electronics. *IEEE Pervasive computing*, *4*(1), pp. 18-27.

Park, C. H., & Kim, Y. G. (2003). Identifying key factors affecting consumer purchase behavior in an online shopping context. *International Journal of Retail & Distribution Management*, *31*(1), pp. 16-29.

Pering, T., Agarwal, Y., Gupta, R., & Want, R. (2006, June). Coolspots: reducing the power consumption of wireless mobile devices with multiple radio interfaces. In *Proceedings of the 4th international conference on Mobile systems, applications and services* (pp. 220-232). ACM.

Poushter, J. (2016, February 22). Smartphone Ownership and Internet Usage Continues to Climb in Emerging Economies. Retrieved from http://www.pewglobal.org/2016/02/22/smartphone-ownership-and-internet-usage-continues-to-climb-in-emerging-economies/.

Rayport, J. & Jaworski, B. (2001). *Introduction to E-Commerce*. New York: McGraw-Hill.

Robins, D., & Holmes, J. (2008). Aesthetics and credibility in web site design. *Information Processing & Management*, *44*(1), pp. 386-399.

Salo, J., & Karjaluoto, H. (2007). A conceptual model of trust in the online environment. *Online Information Review*, *31*(5), pp. 604-621.

Samet, H., Adelfio, M. D., Fruin, B. C., Lieberman, M. D., & Teitler, B. E. (2011, November). Porting a web-based mapping application to a smartphone app. In *Proceedings of the 19th ACM SIGSPATIAL international conference on advances in geographic information systems* (pp. 525-528). ACM.

Sauer, J., & Sonderegger, A. (2009). The influence of prototype fidelity and aesthetics of design in usability tests: Effects on user behaviour, subjective evaluation and emotion. *Applied ergonomics*, *40*(4), pp. 670-677.

Schmidt, K. E., Liu, Y., & Sridharan, S. (2009). Webpage aesthetics, performance and usability: Design variables and their effects. *Ergonomics*,*52*(6), pp. 631-643.

Shankar, V. & Balasubramanian, S. (2009). Mobile Marketing: A Synthesis and Prognosis. *Journal of Interactive Marketing, 23,* pp. 118–129.

Shankar, V., Venkatesh, A., Hofacker, C., & Naik, P. (2010). Mobile marketing in the retailing environment: current insights and future research avenues. *Journal of interactive marketing*, *24*(2), 111-120.

Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, *52*(3/4), 591-611.

Sherif, M., & Hovland, C. I. (1961*). Social judgment: Assimilation and contrast effects in communication and attitude change*. Oxford, England: Yale Univerity Press.

Shneiderman, B. (1998). *Designing the user interface: Strategies for effective user interface interaction*. Reading: Addison Wesley.

Shorey, R., Ananda, A., Chan, M. C., & Ooi, W. T. (2006). *Mobile, wireless, and sensor networks: technology, applications, and future directions*. John Wiley & Sons.

Smith, A. (2014, April 3). Older Adults and Technology Use. Retrieved from http://www.pewinternet.org/2014/04/03/older-adults-and-technology-use/.

Smith, J. (2014, January 29). The Jaw Dropping Effect that Images Can Have on your Conversion Rates. Retrieved from https://www.jeremysaid.com/blog/the-jaw-dropping-effect-that-images-can-have-on-your-conversion-rates/

Sonderegger, A., & Sauer, J. (2010). The influence of design aesthetics in usability testing: Effects on user performance and perceived usability. *Applied ergonomics*, *41*(3), pp. 403-410.

Sridhar, S., & Srinivasan, R. (2012). Social influence effects in online product ratings. *Journal of Marketing*, *76*(5), pp. 70-88.

Stafford, T. F., & Gillenson, M. L. (2003). Mobile commerce: what it is and what it could be. *Communications of the ACM*, *46*(12), pp. 33-34.

Statista (2014a). Number of smartphone users worldwide from 2014 to 2019 (in millions). Retrieved from http://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide/

Straub, D.W. (1994). The Effect of Culture on IT Diffusion: E-Mail and FAX in Japan and the U.S. *Information Systems Research, vol. 5, no. 1,* pp. 23-47

Straub, D.W., Keil, M., & Brenner, W. (1997). Testing the technology acceptance model across cultures: A three country study. *Information & Management, vol. 33,* pp. 1-11

Tarasewich, P. (2003a). Designing mobile commerce applications. *Communications of the ACM*, *46*(12), 57-60.

Tarasewich, P. (2003b). Wireless devices for mobile commerce: user interface design and usability. *Mobile commerce: technology, theory, and applications*, pp. 26-50.

Tarasewich, P., Nickerson, R. C., & Warkentin, M. (2002). Issues in Mobile E-Commerce. *Communications of the Association for Information Systems*: *Vol. 8*, Article 3.

Tractinsky, N., Katz, A. S., & Ikar, D. (2000). What is beautiful is usable.*Interacting with computers*, *13*(2), pp. 127-145.

Tshabalala, S. (2015, July 13). Africa's smartphone market is on the rise as affordable handsets spur growth. Retrieved from http://qz.com/451844/africas-smartphone-market-is-on-the-rise-as-affordable-handsets-spur-growth/

Tuch, A. N., Bargas-Avila, J. A., & Opwis, K. (2010). Symmetry and aesthetics in website design: It'sa man's business. *Computers in Human Behavior*, *26*(6), pp. 1831-1837.

Tversky, A., & Kahneman, D. (1975). Judgment under uncertainty: Heuristics and biases. In *Utility, probability, and human decision making* (pp. 141-162). Springer Netherlands.

Urbany, J. E., Bearden, W. O., & Weilbaker, D. C. (1988). The effect of plausible and exaggerated reference prices on consumer perceptions and price search. *Journal of Consumer research*, *15*(1), 95-110.

van Schaik, P., & Ling, J. (2009). The role of context in perceptions of the aesthetics of web pages over time. *International Journal of Human-Computer Studies*, *67*(1), pp. 79-89.

Varshney, U., & Jain, R. (2001). Issues in emerging 4G wireless networks.*Computer*, *34*(6), pp. 94-96.

Varshney, U., & Vetter, R. (2000). Emerging mobile and wireless networks.*Communications of the ACM*, *43*(6), pp. 73-81.

Venkatesh, V., Ramesh, V., & Massey, A. P. (2003). Understanding usability in mobile commerce. *Communications of the ACM*, *46*(12), pp. 53-56.

Waltz, C. F., Strickland, O., & Lenz, E. R. (1991). *Measurement in nursing research*. Philadelphia: FA Davis Company.

Wang, Y. J., Hernandez, M. D., & Minor, M. S. (2010). Web aesthetics effects on perceived online service quality and satisfaction in an e-tail environment: The moderating role of purchase task. *Journal of Business Research*, *63*(9), pp. 935-942.

Wanous, J. P., Reichers, A. E., & Hudy, M. J. (1997). Overall job satisfaction: how good are single-item measures?. *Journal of applied Psychology*, *82*(2), pp. 247-252.

Webcredible. (2012). Mobile Shopping Behaviour: Key drivers and barriers affecting the adoption of smartphones & mCommerce.

Wilson, G. D. (1966). Arousal Properties of Red Versus Green. *Perceptual and motor skills*, *23,* pp. pp. 942-949.

Wu, C. S., Cheng, F. F., & Yen, D. C. (2008). The atmospheric factors of online storefront environment design: An empirical experiment in Taiwan. *Information & Management*, *45*(7), pp. 493-498.

Wu, S. (2015). Forrester Research Mobile And Tablet Commerce Forecast, 2015 To 2020 (US).

Youngblut, J. M., & Casper, G. R. (1993). Focus on psychometrics single-item indicators in nursing research. *Research in nursing & health*, *16*(6), pp. 459-465.

Yueh, L. (2014, September 24). Is the smartphone market approaching maturity?. Retrieved from http://www.bbc.com/news/business-29339251

Zaczkiewicz, A. (2016, March 7). Amazon, Wal-Mart Lead Top 25 E-commerce Retail List. Retrieved from http://wwd.com/business-news/financial/amazon-walmart-top-ecommerce-retailers-10383750/

Appendix

Appendix A: Figures and Tables

Figure 1: Coversion Rates Smartphone Channels: US vs. Japan

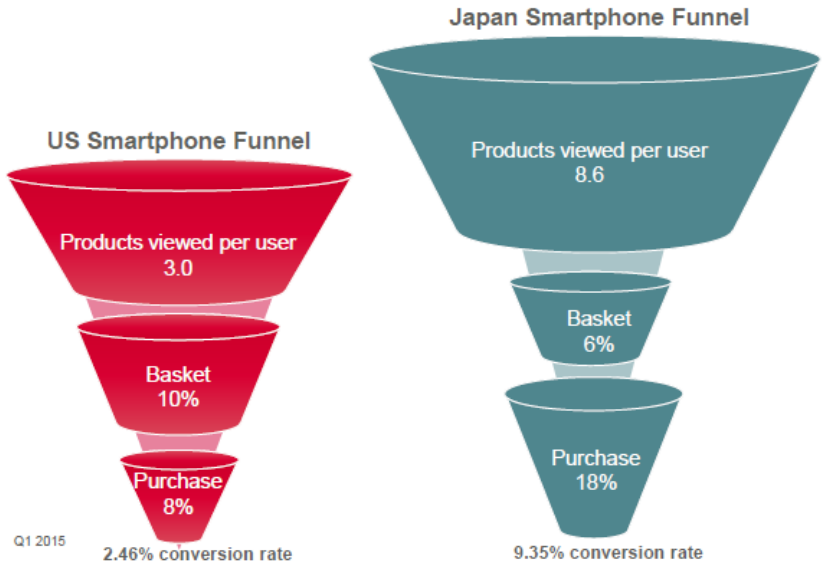


Table 1: Design of Experiment: Screen Attributes Manipulations

Screen	Screen Density	Navigation Design	Symmetry	Color Design	Font Design	Professionalism	Product Rating	Reference Price
1	High	Advanced	Asymmetrical	Cool	Variety	Professional	No Ratings	No RefPrice
2	Low	Advanced	Asymmetrical	Neutral	Variety	Professional	Ratings	No RefPrice
3	Low	Simple	Asymmetrical	Neutral	No Variety	Playful	No Ratings	No RefPrice
4	Low	Simple	Symmetrical	Cool	Variety	Playful	Ratings	No RefPrice
5	Low	Simple	Symmetrical	Cool	No Variety	Professional	Ratings	RefPrice
6	High	Advanced	Symmetrical	Neutral	No Variety	Professional	Ratings	No RefPrice
7	Low	Simple	Asymmetrical	Neutral	Variety	Professional	No Ratings	RefPrice
8	High	Simple	Symmetrical	Neutral	Variety	Playful	No Ratings	No RefPrice
9	High	Simple	Asymmetrical	Cool	Variety	Professional	Ratings	RefPrice
10	Low	Advanced	Symmetrical	Cool	Variety	Playful	No Ratings	RefPrice
11	Low	Advanced	Symmetrical	Cool	No Variety	Professional	No Ratings	No RefPrice
12	High	Simple	Symmetrical	Neutral	No Variety	Professional	No Ratings	RefPrice
13	Low	Advanced	Asymmetrical	Neutral	No Variety	Playful	Ratings	RefPrice
14	High	Advanced	Asymmetrical	Cool	No Variety	Playful	No Ratings	RefPrice
15	High	Simple	Asymmetrical	Cool	No Variety	Playful	Ratings	No RefPrice
16	High	Advanced	Symmetrical	Neutral	Variety	Playful	Ratings	RefPrice

Figure 2: Screen Designs: Rakuten vs. Amazon

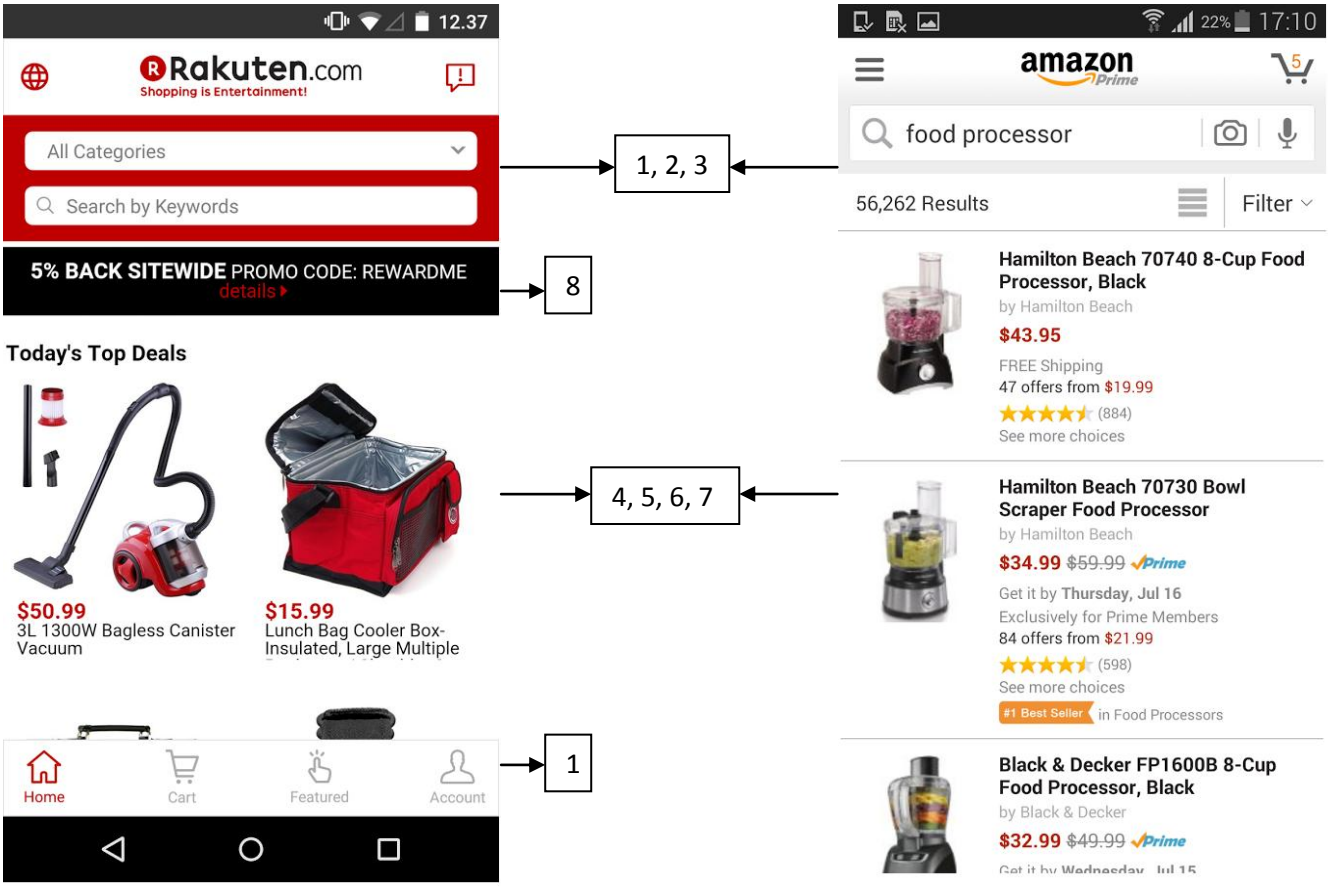
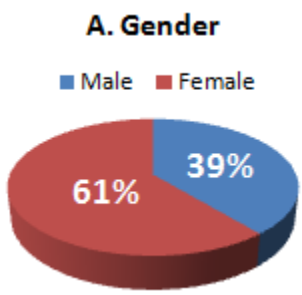
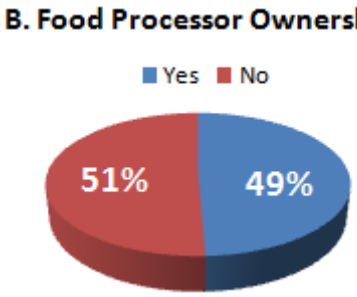
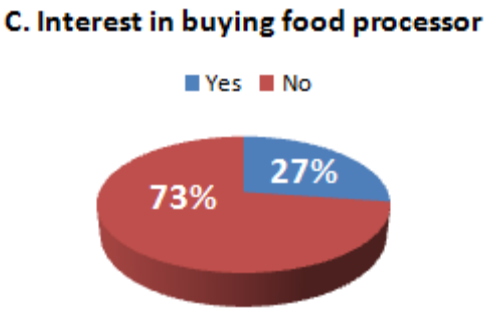
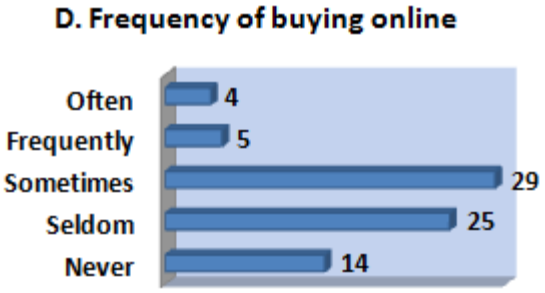


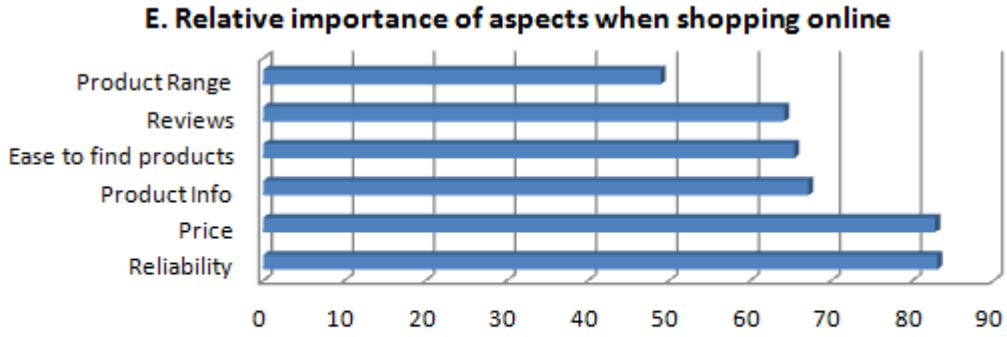
Figure 3: Study Sample: Key Characteristics









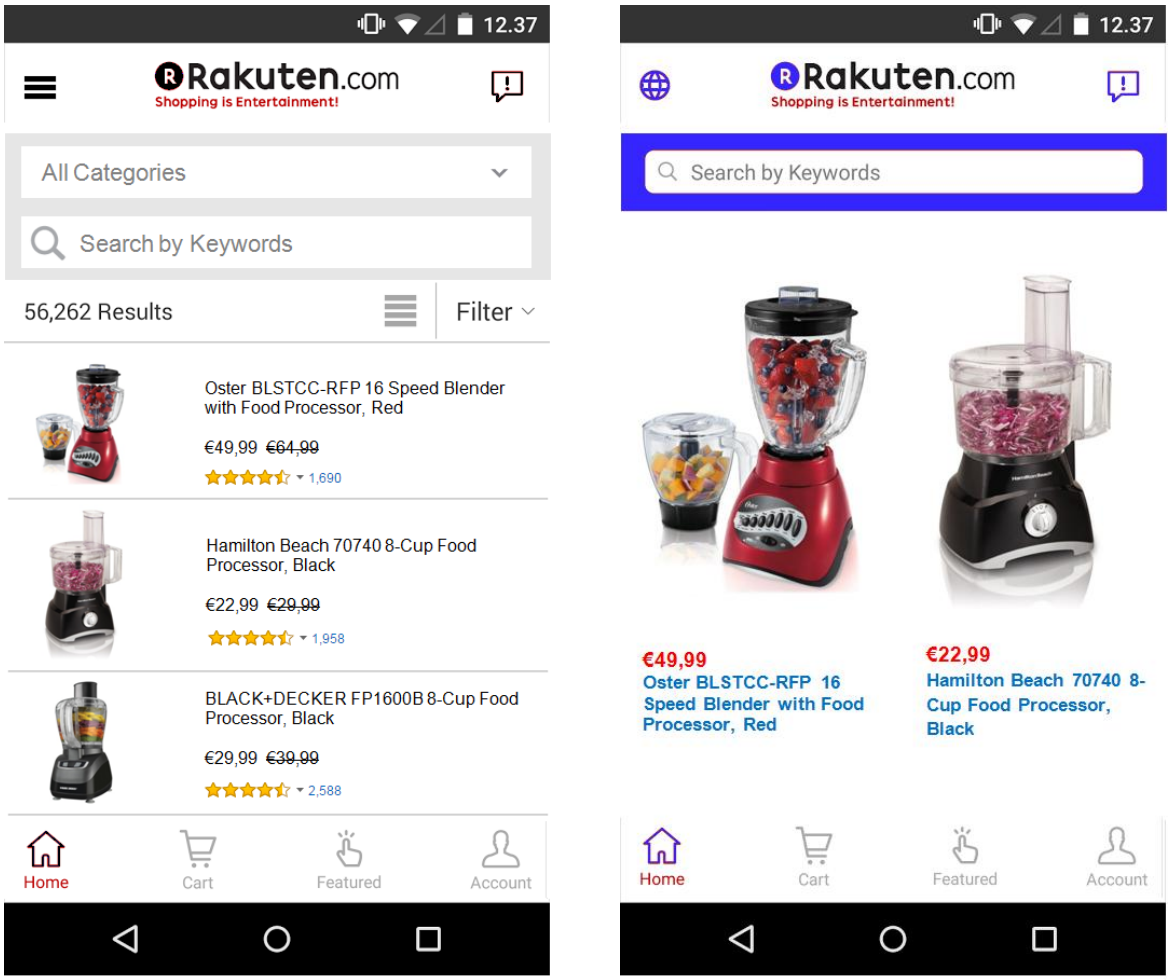


	Reliability	Price	Product Info	Ease to find products	Reviews	Product Range
Mean	83,08	82,82	67,17	65,45	64,23	49,00

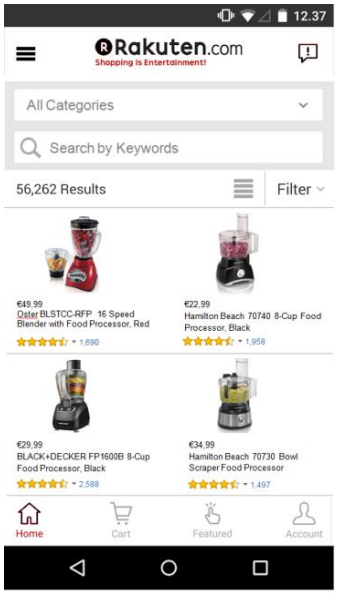
Table 2: Optimal designs PU and WtB: Attribute Specification

	Perceived Usability	Willingness to Buy
Screen Density	High	High
Navigation Design	Simple	Advanced
Symmetry	Asymmetrical	Asymmetrical
Color Design	Neutral	Neutral
Font Design	Variety	No Variety
Professionalism	Professional	Professional
Product Rating	Product Rating	Product Rating
Reference Price	No Reference Price	Reference Price

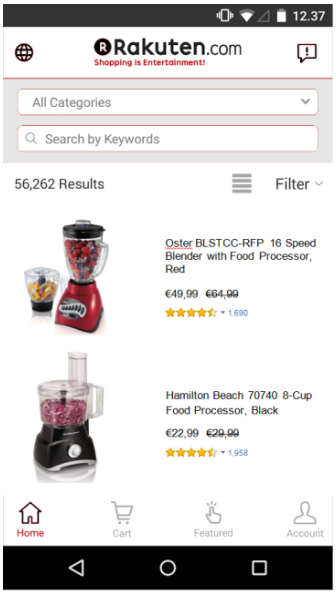
Figure 4: Most Optimal (left) vs. Least Optimal (right)



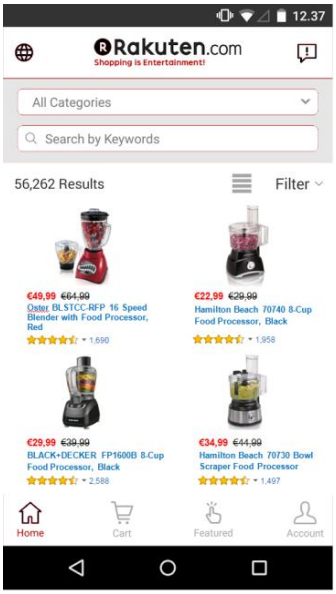
Appendix B: Overall Screen Ratings (1-8)



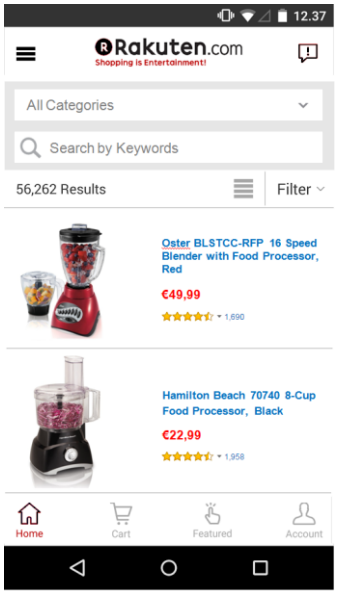
64,82



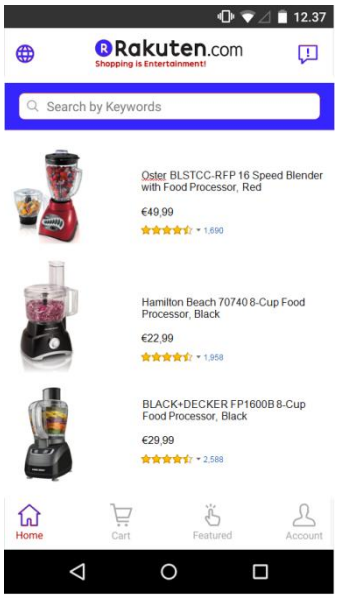
63,05



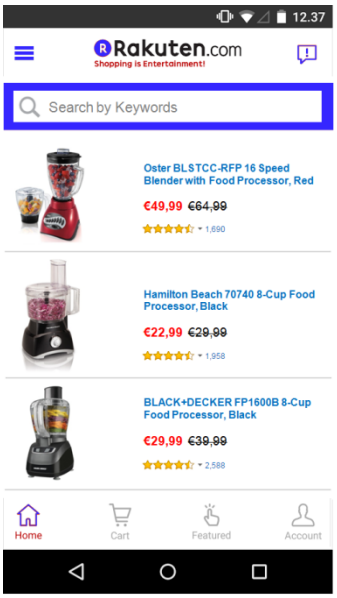
62,82



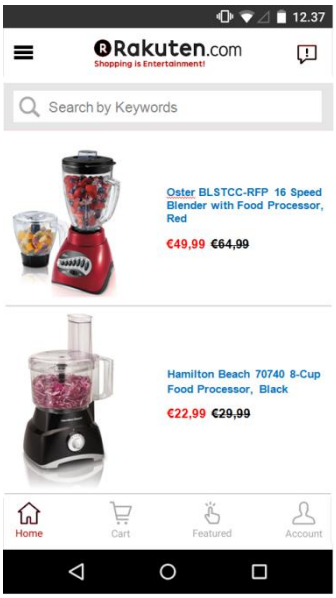
62,65



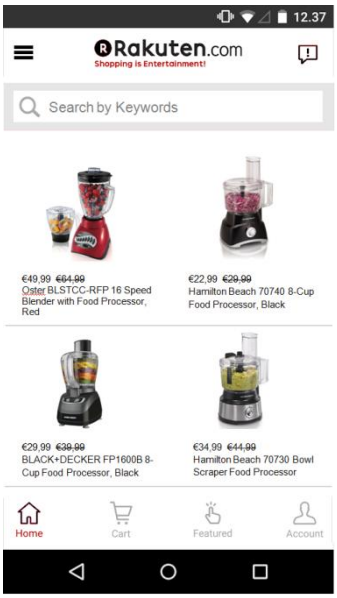
60,82



60,80



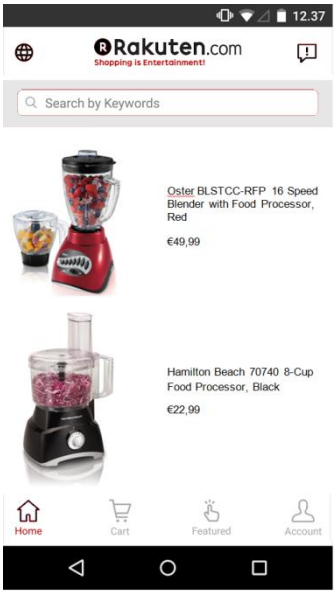
59,54



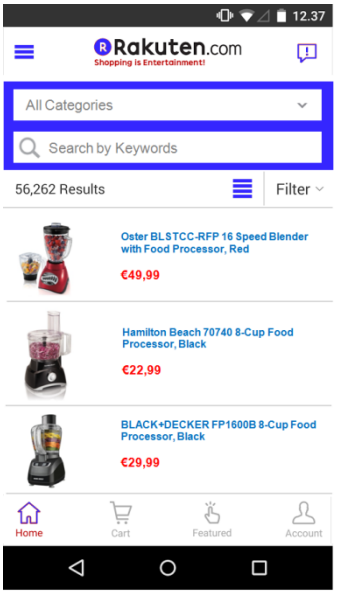
59,06

Overall Screen Ratings (9-16)

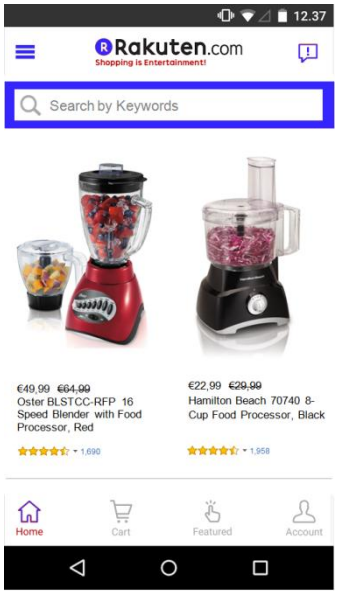
58,19



57,06

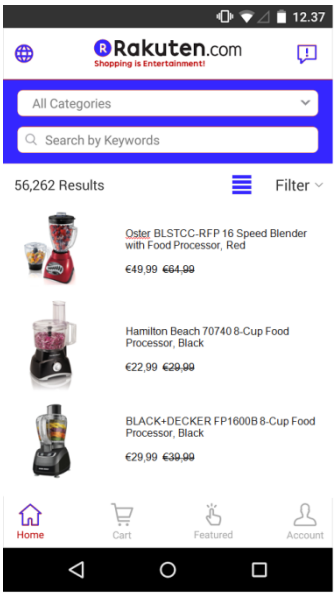


56,90

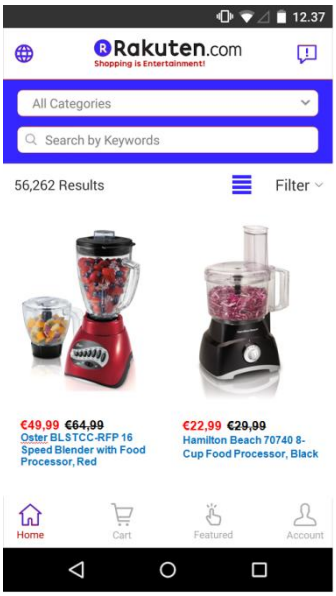


56,77

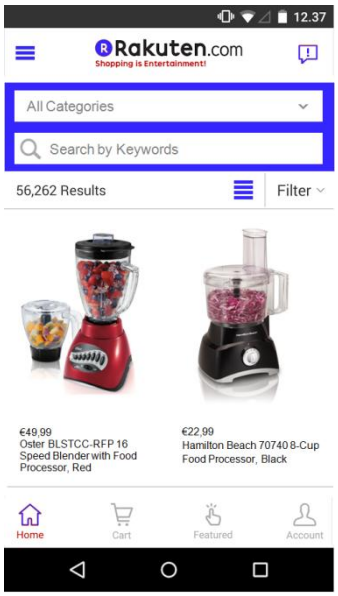
56,28



55,84



54,45



53,91

Appendix C: SPSS Output

Table 1: Respondents with highest SD

ID	Mean	SD[1]	Variance
18	64,81	32,320	1044,563
48	48,81	30,080	904,829
3	49,06	28,163	793,129
75	49,94	26,406	697,263
23	33,94	23,800	566,463
46	63,25	22,773	518,600
36	69,88	22,680	514,383
11	59,25	22,517	507,000
41	54,69	22,045	485,963
20	62,31	21,409	458,363
14	51,69	21,001	441,029
19	50,00	20,412	416,667
53	64,31	19,972	398,896
10	54,38	16,824	283,050
38	70,94	16,454	270,729

1) [1] = Overall SD of PU is 17,087

Figure 1: Deletion of outliers and new scatterplot (1)

Casewise Diagnostics[a]

Case Number	Std. Residual	Perceived Usability	Predicted Value	Residual
215	-3,062	22	70,72	-48,716
216	-3,360	18	71,47	-53,466
273	-3,163	17	67,34	-50,337
755	-3,485	10	65,45	-55,450
760	-3,577	10	66,92	-56,919
764	-3,570	10	66,80	-56,803
767	-3,554	10	66,55	-56,547
1193	-3,137	15	64,92	-49,923

a. Dependent Variable: Perceived Usability

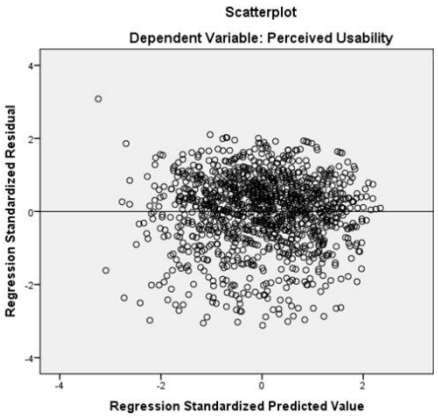


Figure 2: Deletion of outliers and new scatterplot (2)

Casewise Diagnostics[a]

Case Number	Std. Residual	Perceived Usability	Predicted Value	Residual
42	3,081	90	42,77	47,225
354	-3,057	10	56,86	-46,863
358	-3,031	15	61,46	-46,463
365	-3,119	17	64,82	-47,821
569	-3,014	22	68,20	-46,200

a. Dependent Variable: Perceived Usability

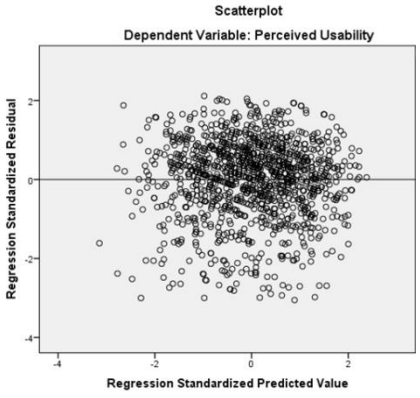


Table 2: Extended Model (PU)

Coefficients[a]

Extended Model**		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	45,301	4,837		9,366	,000*
	Screen Density	2,805	,900	,085	3,115	,002*
	Navigation Design	-,018	,901	-,001	-,020	,984
	Symmetry	-1,335	,900	-,041	-1,482	,139
	Color Design	-4,244	,901	-,129	-4,713	,000*
	Font Design	,504	,900	,015	,560	,575
	Professionalism	1,039	,901	,032	1,154	,249
	Product Rating	3,899	,901	,119	4,329	,000*
	Reference Price	-,485	,901	-,015	-,539	,590
	Gender	-,585	,975	-,017	-,600	,548
	Frequency of buying online	1,023	,475	,064	2,153	,031*
	Impact of Price	,072	,044	,049	1,655	,098
	Impact of Reviews	,049	,028	,060	1,714	,087
	Impact of Ease to find products	-,002	,027	-,002	-,062	,951
	Impact of Reliability	-,017	,032	-,018	-,541	,589
	Impact of Product Range	,141	,031	,184	4,623	,000*
	Impact of Product Info	,019	,038	,019	,489	,625
	Food Processor Ownership	1,842	,955	,056	1,928	,054
	Interest in buying a Food Processor	-1,664	1,055	-,045	-1,578	,115

a. Dependent Variable: Perceived Usability

1) * = Statistically significant at 1%-level

2) ** = R Square of model: 0,097, ANOVA: F: 7,197 -> p = 0,000*

Table 3: Testing for significance of interaction variables

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	47,109	5,153		9,142	,000
	Screen Density	3,684	1,158	,112	3,182	,002
	Navigation Design	-,032	1,158	-,001	-,028	,978
	Symmetry	-1,252	1,158	-,038	-1,081	,280
	Color Design	-5,506	1,158	-,168	-4,755	,000
	Font Design	,381	1,158	,012	,329	,742
	Professionalism	1,243	1,158	,038	1,073	,284
	Product Rating	3,757	1,158	,114	3,244	,001
	Reference Price	,165	1,158	,005	,143	,887
	Gender	-4,036	3,723	-,120	-1,084	,279
	Frequency of buying online	,407	,642	,026	,634	,526
	Impact of Price	,072	,045	,049	1,615	,107
	Impact of Reviews	,052	,029	,065	1,804	,071
	Impact of Ease to find products	-,004	,028	-,005	-,147	,883
	Impact of Reliability	-,016	,033	-,016	-,499	,618
	Impact of Product Range	,148	,032	,193	4,663	,000
	Impact of Product Info	,010	,040	,010	,245	,807
	Food Processor Ownership	1,807	1,254	,055	1,441	,150
	Interest in buying a Food Processor	-2,352	1,338	-,064	-1,757	,079
	MaleSD	-2,230	1,845	-,054	-1,209	,227
	MaleND	,011	1,845	,000	,006	,995
	MaleSym	-,185	1,845	-,004	-,100	,920
	MaleCD	3,202	1,845	,078	1,735	,083
	MaleFD	,315	1,845	,008	,171	,865
	MaleProf	-,522	1,845	-,013	-,283	,777
	MalePR	,347	1,845	,008	,188	,851
	MaleRP	-1,653	1,845	-,040	-,896	,371
	MaleFreqBO	1,373	,920	,112	1,494	,136
	MaleOwnFP	-,136	2,029	-,003	-,067	,947
	MaleIntFP	1,851	2,192	,033	,844	,399

a. Dependent Variable: Perceived Usability

Table 4: Multicollinearity testing

		SD	ND	Sym	CD	FD	Prof	PR	RP	Gender	FreOn	IPrice
PU	**Tolerance**	1,000	1,000	,606	,606	1,000	1,000	1,000	1,000	,320	,845	,839
	VIF	1,000	1,000	1,650	1,650	1,000	1,000	1,000	1,000	3,128	1,184	1,192
		IRev	**IEase**	**IRel**	**IPrRange**	**IPrInfo**	**FPOwn**	**FPBuy**	**MaleCD**	**OwnBuy**	**MaleSym**	**MaleOFP**
	Tolerance	,614	,683	,682	,475	,499	,637	,523	,377	,467	,377	
	VIF	1,629	1,465	1,466	2,105	2,002	1,569	1,912	2,650	2,141	2,650	
		SD	**ND**	**Sym**	**CD**	**FD**	**Prof**	**PR**	**RP**	**Gender**	**FreOn**	**IPrice**
WtB	**Tolerance**	1,000	1,000	1,000	,610	1,000	1,000	1,000	1,000	,330	,848	,809
	VIF	1,000	1,000	1,000	1,638	1,000	1,000	1,000	1,000	3,034	1,179	1,236
		Irev	**IEase**	**IRel**	**IPrRange**	**IPrInfo**	**FPOwn**	**FPBuy**	**MaleCD**	**OwnBuy**	**MaleSym**	**MaleOFP**
	Tolerance	,598	,654	,686	,455	,473	,414	,509	,379	,466		,381
	VIF	1,673	1,529	1,458	2,198	2,116	2,418	1,966	2,638	2,148		2,627

Table 5: Added interaction variables (PU)

	Unstandardized Coefficients		Standardized Coefficients		
Variable	B	Std. Error	Beta	t	Sig.
MaleSym	-,181	1,846	-,004	-,098	,922
MaleCD	3,200	1,846	,078	1,733	,083
OwnFP_BuyFP	-21,474	2,068	-,400	-10,384	,000*
Int_Sym	3,046	1,942	,063	1,569	,117
Int_CD	-2,976	1,942	-,062	-1,533	,126

1) * = Statistically significant at 1%-level

Table 6: Added interaction variables (WtB)

	Unstandardized Coefficients		Standardized Coefficients		
Variable	B	Std. Error	Beta	t	Sig.
MaleCD	4,024	2,057	,083	1,956	,051
MaleOwnFP	-17,701	2,241	-,333	-7,900	,000*
OwnFP_BuyFP	-15,521	2,423	-,246	-6,405	,000*
Int_CD	-3,490	2,308	-,062	-1,512	,131

1) * = Statistically significant at 1%-level

Table 7: Full Model (PU)

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	45,797	4,672		9,803	,000
	Screen Density	2,841	,861	,087	3,298	,001
	Navigation Design	-,073	,861	-,002	-,084	,933
	Symmetry	-2,211	1,247	-,067	-1,773	,076
	Color Design	-4,761	1,247	-,145	-3,818	,000
	Font Design	,488	,861	,015	,567	,571
	Professionalism	1,058	,861	,032	1,228	,220
	Product Rating	3,854	,861	,117	4,474	,000
	Reference Price	-,527	,861	-,016	-,611	,541
	Gender	-1,044	1,561	-,031	-,668	,504
	Frequency of buying online	,918	,455	,058	2,018	,044
	Impact of Price	,099	,042	,067	2,355	,019
	Impact of Reviews	,055	,027	,068	2,022	,043
	Impact of Ease to find products	-,024	,026	-,030	-,948	,343
	Impact of Reliability	-,085	,031	-,086	-2,696	,007
	Impact of Product Range	,153	,029	,199	5,240	,000
	Impact of Product Info	,031	,037	,031	,834	,405
	Food Processor Ownership	7,816	1,079	,238	7,244	,000
	Interest in buying a Food Processor	7,369	1,919	,200	3,840	,000
	MaleSym	,061	1,767	,001	,034	,973
	MaleCD	3,101	1,767	,075	1,755	,080
	OwnFP_BuyFP	-21,504	2,063	-,400	-10,424	,000
	Int_Sym	3,057	1,942	,064	1,574	,116
	Int_CD	-2,762	1,942	-,058	-1,422	,155

a. Dependent Variable: Perceived Usability

Table 8: Hierarchical Model (WtB)

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	52,057	1,632		31,907	,000
	Screen Density	1,646	1,088	,043	1,513	,130
	Navigation Design	,756	1,088	,020	,696	,487
	Symmetry	-1,315	1,088	-,034	-1,209	,227
	Color Design	-3,831	1,088	-,099	-3,522	,000
	Font Design	-,344	1,088	-,009	-,316	,752
	Professionalism	,627	1,088	,016	,576	,565
	Product Rating	4,276	1,088	,111	3,931	,000
	Reference Price	,981	1,088	,025	,901	,368
2	(Constant)	44,933	5,518		8,144	,000
	Screen Density	1,646	1,027	,043	1,603	,109
	Navigation Design	,756	1,027	,020	,737	,461
	Symmetry	-1,315	1,027	-,034	-1,281	,201
	Color Design	-3,831	1,027	-,099	-3,732	,000
	Font Design	-,344	1,027	-,009	-,335	,738
	Professionalism	,627	1,027	,016	,610	,542
	Product Rating	4,276	1,027	,111	4,165	,000
	Reference Price	,981	1,027	,025	,955	,340
	Gender	-4,750	1,114	-,120	-4,265	,000
	Frequency of buying online	,099	,542	,005	,183	,855
	Impact of Price	-,059	,050	-,034	-1,187	,236
	Impact of Reviews	,020	,032	,021	,629	,530
	Impact of Ease to find products	,094	,031	,098	3,042	,002
	Impact of Reliability	-,123	,037	-,106	-3,363	,001
	Impact of Product Range	,072	,035	,080	2,077	,038
	Impact of Product Info	,130	,044	,112	2,984	,003
	Food Processor Ownership	4,625	1,090	,120	4,244	,000
	Interest in buying a Food Processor	6,768	1,201	,156	5,635	,000

a. Dependent Variable: Willingness to Buy

Table 9: Full Model (WtB)

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	50,881	5,327		9,552	,000
	Screen Density	1,646	,979	,043	1,682	,093
	Navigation Design	,756	,979	,020	,773	,440
	Symmetry	-1,315	,979	-,034	-1,344	,179
	Color Design	-4,447	1,414	-,115	-3,145	,002
	Font Design	-,344	,979	-,009	-,352	,725
	Professionalism	,627	,979	,016	,640	,522
	Product Rating	4,276	,979	,111	4,369	,000
	Reference Price	,981	,979	,025	1,002	,317
	Gender	2,451	1,749	,062	1,402	,161
	Frequency of buying online	-,105	,517	-,006	-,203	,839
	Impact of Price	-,116	,048	-,067	-2,391	,017
	Impact of Reviews	,073	,031	,077	2,343	,019
	Impact of Ease to find products	,023	,030	,024	,757	,449
	Impact of Reliability	-,178	,036	-,153	-4,993	,000
	Impact of Product Range	,142	,034	,157	4,183	,000
	Impact of Product Info	,051	,043	,044	1,199	,231
	Food Processor Ownership	16,798	1,522	,436	11,038	,000
	Interest in buying a Food Processor	17,295	1,894	,400	9,132	,000
	MaleCD	3,818	2,012	,078	1,898	,058
	MaleOwnFP	-19,058	2,187	-,359	-8,715	,000
	OwnFP_BuyFP	-17,197	2,350	-,272	-7,317	,000
	Int_CD	-3,194	2,203	-,057	-1,450	,147

a. Dependent Variable: Willingness to Buy

Appendix D: Literature review Technology, Security and Privacy

1 Technology

If it wasn't for technology, we would still live in a primitive world. Since the introduction of the first personal computer, digital technology has increased rapidly. As mentioned before, technology has the potential to solve most of the issues that are currently being faced in m-commerce. In paragraph 2.2, several technological solutions has already been discussed that can solve issues related to usability. However, this paragraph revolves around limitations in the current technology. Whereas issues with usability are often directly experienced by users, consumers are not always aware of the technological limitations that could hinder them in using a mobile device. However, when these issues reach a critical point, the impact is much bigger than other issues. Think about an empty battery, an application that can't be accessed or a slow internet connection. In all cases a customer won't be able to engage in m-commerce activities. Therefore it is important to thoroughly investigate these issues and look into possible solutions.

In the early stage, new technologies have to face a considerable amount of technological challenges to ultimately be adopted by a broad audience. In these early years, there is typically a lot of research on the new technology and how to improve it. In the case of mobile technology, this was not different. Although some lessons could be learned from the wired counterpart, mobile technology has proven to be very different (Coursaris & Hassanein, 2002; Dholakia & Dholakia, 2004). Consequently, scientists researched a broad variety of technological issues that were trending in the early stage of mobile technology. Topics of research included: Analysis of second generation (2G) and 2,5G networks and technologies and predicting the possibilities of 3G and 4G networks (Buellingen & Woerter, 2004; Coursaris, & Hassanein, 2002; Shorey et al, 2006; Varshney & Jain, 2001; Varshney & Vetter, 2000), analysis of WiFi networks (Al-Alawi, 2006; Lehr & McKnight, 2003), wireless and mobile protocols (Bannan, 2000; Barnes, 2002; Juul & Jorgensen, 2003), power management (Paradiso & Starner, 2005; Pering et al, 2006) and research on various possible applications, including mobile banking (Herzberg, 2003), mobile entertainment (McInness et al, 2002; Salo & Karjaluoto, 2007) and mobile ticketing (Bauer et al, 2007).

However, the early stage of mobile technology was over a decade ago. Since then the mobile technology has made some remarkable progress. Smartphones has more or less become small computers with a broad range of functionalities. Smartphone technology has become standardized, which have lead to lower prices. Manufacturers are targeting the budget end of the market, as smartphone markets in developed countries have matured (Yueh, 2014). Due to the considerable progress in the mobile technology and the subsequent standardization, technological issues haven't been much of a hot topic lately. However, this doesn't mean that the current technology is without any flaws or limitations. These flaws and limitations are discussed below.

The backbone of mobile technology are the wireless networks. These networks permit users to transmit data between mobile and other computing devices without the requirement of a wired connection (Coursaris & Hassanein, 2002). As there are many different wireless network systems, they are often being categorized in so called 'generations', based on the period in which they were invented. In the early stage, the speed of data transmission was low and network coverage was limited. However, with every update these parts are significantly improved, along with newly introduced opportunities for more functionalities and applications. Still, there's always demand for bigger and faster data transmission, more advanced functionalities and better coverage. While 4G is currently still paving its way in the smartphone market, fifth generation (5G) networks are already being developed (Latha et al, 2014).

A study by Webcredible (2012) shows that improved networks are indeed a necessity. The study identified slow and unstable connections as one of the main issues that concerned smartphone users. Their fear to be cut off in the middle of a mobile transaction makes them prefer buying at home. Back in 2002, Coursaris & Hassanein already identified this as one of the main connectivity related issues. However, in that time third generation networks were not even fully developed, implying that data transmission was considerably slower. So how can the findings of Webcredible's study, which was conducted at the time 4G networks were introduced, be explained?

One factor that could explain this is the fact that networks in reality never achieve the peak download speed that it should be capable of (Cha, 2015). This implies that connections could get unstable and slow, whilst the network in theory should be capable of achieving speeds that makes the connection works smoothly. Another factor that could explain connectivity issues is coverage. Although most western countries have almost full coverage, there are still areas in which the coverage is weak. Network infrastructures are typically laid around cities, in which the connection is strong. However, rural areas could experience problems with coverage as the network infrastructure isn't as strong compared to urban areas. Looking at a global scale, there are still some major coverage gaps to be filled. Ten percent of the global population doesn't have access to basic text and voice services, while thirty percent of the population lacks access to the internet at 3G speeds (GSMA Intelligence, 2015). The countries and areas in which the coverage is limited are typically thinly populated areas with low purchasing power. The high fixed costs of laying and maintaining a network infrastructure often doesn't outweigh the potential benefits, making it uneconomical to invest in networks in these areas.

Another issue related to wireless networks, pointed out by Coursaris & Hassanein (2002), is the costs of having access to a mobile network. Users can subscribe to different transfer rates, which are differently priced. Needless to say, subscriptions with fast data speeds comes with higher costs. For some customers this could not be affordable, leaving them with

slower data speed. That will lead us back to the first issue, in which users prefer to buy online via desktops due to slow and unstable wireless connections.

Next to network related issues, the functioning of mobile devices itself could also hinder the user in engaging in m-commerce activities. The two major issues with mobile devices are screen size limitations and battery lifetime. In chapter 2.2.1, screen size limitations has already been discussed extensively. One key takeaway was that there is a trade-off between usability and portability. Smaller screens imply higher portability but lower usability, while larger screens imply higher usability but lower portability, also due to higher energy consumption. Efficient energy consumption is an important subject these days. Mobile devices are getting increasingly advanced, with larger screens and increased screen resolution. Consequently, energy consumption has significantly increased. Taking into consideration that devices will shrink in size again and devices will get even more advanced, batteries should be reduced in size and at the same time have improved battery lifetime to keep up with the newest technologies (Mousa, 2012).

It's evident that the current mobile technology still has its flaws and needs further improvements to establish a financially accessible, fast and stable wireless network around the globe. Although all the technological issues can't be fixed by just one solution, most of the abovementioned issues actually can be fixed by one solution: 5G networks. Fifth Generation networks are still under development and will probably take some time before they are introduced, but several studies have already studied the potential benefits of 5G networks (Hema Latha et al, 2014; Khan, 2014; Mousa, 2012). Their main findings are discussed below:

- With 5G networks, data speeds will significantly increase. Downloads can be completed over 50 times as fast as 4G networks (Cha, 2015). This will ensure fast and stable connections, making it more attractive for users to engage in m-commerce activities.
- There will be a lower outage (loss of signal) probability, which implies higher data rates available and better signal coverage at the edges of cell areas, increasing overall coverage.
- Latency, the time it takes to send a packet of data from one to another device, will be significantly decreased. Less time delay means that applications will work smoother and advanced technologies (e.g. driverless cars) get more and more likely to be introduced.
- Fifth Generation networks will be able to interconnect most of the already existing communication infrastructures. Consequently, 5G has the capability and capacity to create a network that can accommodate billions of connected devices.
- Due to more advanced and energy-effective technology, 5G networks require less energy consumption. This enhances the battery lifetime and portability of devices.

It can be concluded that 5G networks have the features that could solve almost every issue related to connectivity. In addition to that, energy consumption can be made more efficiently by 5G networks. However, fifth generation networks aren't capable of directly fixing the cost issue. Although studies suggest that traffic fees will get cheaper due to lower

infrastructure deployment costs (Kahn, 2014; Mousa, 2014) and cost per bit will be low (Hema Latha et al, 2014), it is questionable if 5G networks will become financially accessible for the entire global population. According to Goldman (2015), wireless companies would have to reduce the price of each bit of data to 1/1000 of what they cost today if you want to have 5G service to cost exactly the same as 4G service nowadays. Therefore, wireless companies should even put more focus on reducing costs, so that they can offer data at a significantly lower price. If companies manage to achieve this, 5G will undoubtedly be the future of mobile technology.

2 Security and Privacy

Ever since the introduction of mobile technology, security and privacy related issues has been hot topics. Several studies in the early years pointed out the importance of solving these issues. These studies extensively discussed the various security and privacy risks in m-commerce (Ghosh & Swaminatha, 2001) and recognized the need for a unified security standard (Barnes, 2002; Tarasewich, Nickerson & Warkentin, 2002) and data protection (Buellingen & Woerter, 2004). One would think that a decade later, with many technological improvements in the field of m-commerce, this issue has been reduced to an issue of lower importance. However, a study by research company Webcredible (2012) proved that security issues are still a major issue. According to this study, the biggest issue that prevented smartphone users from engaging in m-commerce activities are security concerns, being even bigger than issues related to usability and connectivity. Participants stated that they were worried about having their phone hacked or infected by viruses, resulting in their personal or card details being intercepted and used for criminal purposes.

The findings of this study are not surprising, since cybercriminals are moving their activities towards mobile devices. One reason for this is the absence of security software on smartphones. In 2013, a mere 17 percent of smartphone users had installed security software on their phones, making them extremely vulnerable for cyber attacks (Curtis, 2013). Furthermore, 25 percent of those surveyed claimed to have been victim of mobile cybercrime within the last 12 months, costing the society close to £1 billion that year in the UK alone. These statistics confirm that security in the mobile environment is a big current issue. Next to the direct costs caused by cybercrime, safety concerns costs companies lots of money. Studies found that perceived safety of a website or application significantly influences trust and intention to use, which ultimately influences customers' willingness to buy (Dai & Palvia, 2009; Flavián and Guinalíu, 2006). This implies that as long as customers have safety concerns and therefore refrain from engaging in mobile commerce activities, companies will experience substantial losses of potential revenue. It is found that this issue is especially important for applications and websites that require high interaction and containing personal information (Coursaris & Hassanein, 2002; Magura, 2003). In case of low involvement websites and applications, security concerns were found to be of very low

importance. Nevertheless, it is important for all parties involved in m-commerce that security risks will be reduced in the upcoming years.

One way to solve this issue is by developing and implementing better security software. In the feature, more advanced technology could make it harder for cybercriminals to perform criminal activities via mobile devices. it is expected that the fifth generation network will be more secure because of better cognitive radio/SDR Security (Khan, 2014; Mousa, 2012), although Mousa also points out that because of the new possibilities of 5G, new security challenges will arise. Furthermore, Tarasewich suggested back in 2003 that biometrics could increase the security of mobile applications. Nowadays, biometric authentication has indeed been introduced, although few devices support this kind of authentication (Niranjanamurthy & Kavyashree, 2013). Therefore, it is important that all new devices support this kind of authentication software. Next to these technological improvements in terms of security, it is important that a unified security standard will be established. Along with the 5G network improvements, an up to date security standard could significantly enhance the security in mobile commerce.

However, merely the technological improvements won't do the job. Increasing security services is one thing, but probably the most important thing is to make your customers aware of the fact that m-commerce is safe (Ivan, Milodin & Zamfirou, 2013). Although security software for mobile devices can still be improved, there is plenty of solid security software available for these devices. The problem though is that customers are not aware of this. In 2013, 54 percent of UK smartphone users wasn't aware of the existence of security solutions for mobile devices (Curtis, 2013). It is clear that a big part in overcoming security concerns in mobile commerce lies in making people aware of the existence of security software and the risks of not having any security installed on their mobile devices. If companies can get customers to trust their websites, applications and payment methods, then mobile commerce will significantly grow in the upcoming years.

Next to security concerns, privacy concerns also prevent users from engaging in m-commerce activities. Mobile commerce provides companies with huge opportunities to collect all kinds of personal data, which could be used to create detailed customer profiles. These profiles on their turn can be used to predict customer preferences and make specific offers to these customers through mobile advertising (Kumar et al, 2013). This so-called data-driven marketing has gained increasing popularity in the past years. The question however, is whether it is ethically accepted or not. Obviously, online shops do need a certain amount of personal information in order to do business with their customers. However, businesses may use this data beyond the original purpose or even collect information about customers of which they are not aware of. That makes this kind of data gathering highly questionable in regards to users' privacy (Cleff, 2007). Customers are often unaware of the amount of information that is being collected and the purpose it is used for (Gosh &

Swaminatha, 2001). This leads to Direct and Indirect Anonymity issues, which hinders smartphone users from engaging in m-commerce activities (Beach, Gartrell & Han, 2009).

The problem with the use of personal information is that it could also be considered as beneficial. Collecting personal data allows companies to send helpful information to customers, which could enhance the overall m-commerce experience as personalized services can satisfy both customers and businesses (Cleff, 2007; Kumar et al., 2013). On the other hand though, excessive collection of personal information can make customers reluctant to share their personal information. One solution that could solve this problem is to give customers more control on to what extent their personal information is used. Shankar and Balasubramanian (2009) proposed that customers should be given the option to opt-in in case that they consider it beneficial to receive specific messages and offers. Cleff (2007) adds to this that the customer shouldn't be encouraged to opt-in into a blanket provision for the use of their personal data, but should be freely and voluntarily be able to choose what type of data is collected about them.

Furthermore, proper legislation should be the basis of privacy protection. Cleff (2007) extensively discussed the requirements in terms of legislation to ensure the privacy of customers in the mobile commerce environment. The core of this legal framework is that businesses are required to inform the user about the applied information practices and that the customer is given the opportunity to choose whether or not to disclose personal data and receive mobile advertising. Not following this legislation should be regarded as an unlawful and punishable act. Next to that, it is important that users of mobile devices receive proper education about the dangers and benefits of disclosing certain personal information. They should be made aware of the fact that they have control over the extent of information that is being used by businesses. If companies can gain their customers' trust in regards to ensuring their privacy, m-commerce can significantly benefit from this.

Appendix E: Normality tests of dependent variables and testing assumptions of multiple regression

1 Normality Tests of Dependent Variables

In this part, normality test are performed to test whether it can be assumed that the dependent variables perceived usability and willingness to buy are normally distributed. This will provide valuable insights about how both variables are distributed.

1.1 Perceived Usability

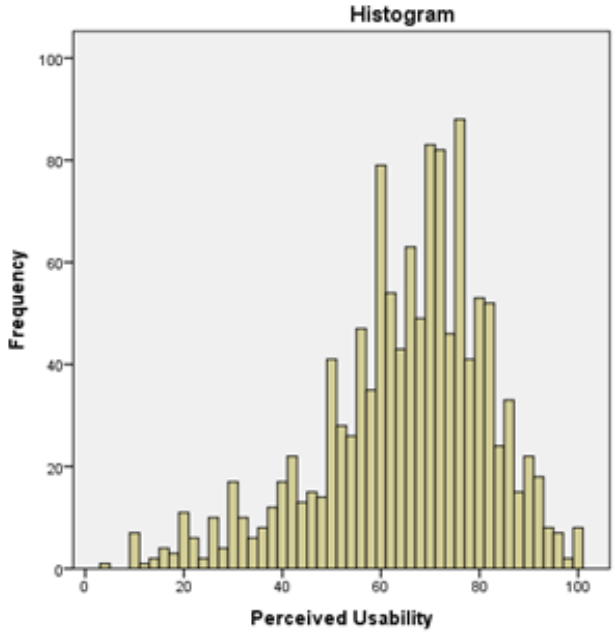
In statistics, there are various ways to test for normal distribution of data. SPSS provides options for several of these tests to be performed within just a couple of clicks. A normality test that is often used, and is standard reported in the SPSS Explore procedure is the Shapiro-Wilk test, developed by scientists Samuel Shapiro and Martin Wilk (1965). This test assumes that the data are not-normally distributed, which means that a p-value lower than 0,05 implies a significant non-normal distribution of the data. Although a second test of normality is reported along with the SW test, the Kolmogorov-Smirnov (KS) test, the Shapiro-Wilk test is widely regarded as the most powerful test of normality (Oztuna et al., 2006).

The p-value of the SW test on perceived usability is 0,000, which implies that PU follows a non-normal distribution. However, according to Elliott and Woodward (2007) this test should only be interpreted in case that the sample size is less than 50. In this case though, the sample size is considerably larger (N = 1.232). Therefore, the SW test statistic will not be interpreted. Instead, graphical methods will be used to reveal a possible normal distribution of the data. According to Oztuna et al. (2006), there are five graphical methods to test normality: Histogram, Stem and Leaf Plot, Boxplot, Normal Quantile Quantile Plot (Q-Q Plot) and Normal Probability Plot (P-P Plot). In this section, we will investigate the Histogram, the Boxplot and the Q-Q Plot. The P-P Plot rather tests for the assumption of normally distributed errors and is therefore not included in this section. However, in the part 2 of this appendix, P-P Plots will be discussed.

Figure 1 on the next page contains a histogram, a boxplot, a Q-Q Plot and additional statistical information about the distribution of the dependent variable perceived usability. Looking at the histogram, it can be noted that the distribution doesn't exactly follow a normal distribution, as the data are negatively skewed (Skewness of -0,781). In literature, there have been many interpretations of the acceptable range of skewness. In this study however, the interpretation of Bulmer (2012) is used, which deems a skewness value between -1 and 1 as acceptable, i.e. not extremely skewed. The second graphical instrument to measure a normal distribution is a boxplot. The boxplot in figure 1 shows a large number of outliers at the bottom of the plot. One-sided outliers suggests a skewed distribution of the data (Oztuna et al., 2006), which was already concluded from the histogram.

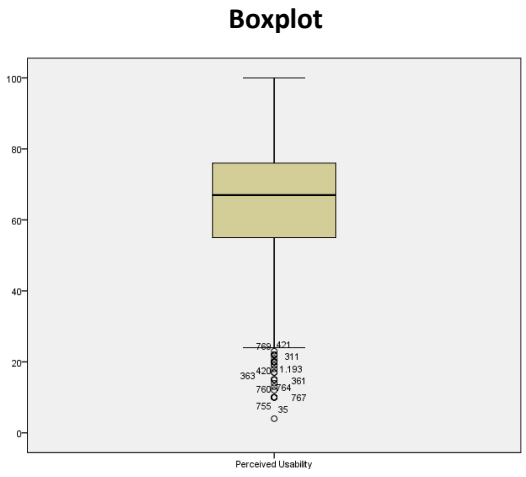
The large amount of outliers can possibly be explained by the unusual high standard deviations of some of the respondents. When comparing the standard deviations of the PU ratings of separate respondents, it can be noted that some respondents had an unusually high standard deviation. Table 1 in Appendix C provides information about the 15 respondents with the highest standard deviations. Taking into account the fact that the overall standard deviation of PU (of 77 respondents) is 17,087, it is interesting to see that only 13 respondents have a SD that is higher than the overall standard deviation of perceived usability, suggesting that the SD of these respondents is unusually high. Although it is unlikely that these values reflects the realistic responses of these respondents, the ratings of those respondents are not deleted as they can still have predictive power in regard to the importance of the screen attributes. However, it is assumed that realistic responses will decrease the negative skewness of the data, resulting in a more normal distribution of PU. Furthermore, the negative skewness can be explained by the fact that the mean of PU lies closer to the maximum possible value of the scale than the minimum possible value.

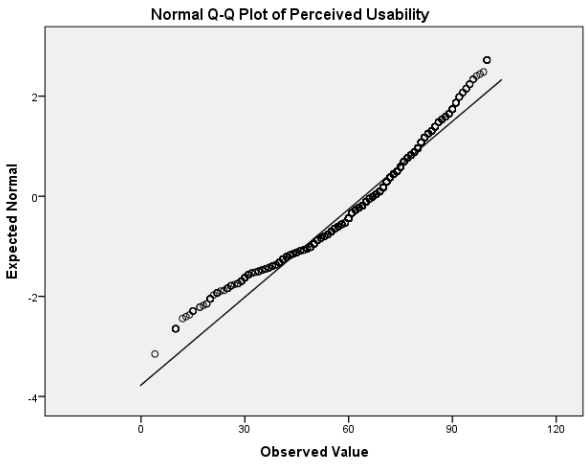
Figure 1: Graphical testing for normal distribution of PU



		Statistic	St. Dev.
Mean		64,42	,487
95% Confidence Interval for Mean	Lower Bound	63,46	
	Upper Bound	65,37	
Median		67,00	
Variance		291,949	
Std. Deviation		17,087	
Minimum		4	
Maximum		100	
Range		96	
Interquartile Range		21	
Skewness		-,781	,070
Kurtosis		,627	,139
Shapiro-Wilk Test		,961*	

1) * = statistically significant at 1%-level





The last graphical tool to test for normality of the distribution is the Q-Q Plot, which is the bottom right plot shown in figure 1. Normal Q-Q plots plot the quantiles of a variable's distribution against the quantiles of the normal distribution. Typically, in the case that data are normally distributed, the points in the Q-Q plot are all lying on or near the straight line drawn through the middle half of the points (Oztuna et al., 2006). The points in the Normal Q-Q Plot of Perceived Usability tend to follow the straight line in the middle. However, at the ends of the straight line, they tend to deviate considerably from the line, which would dispute a normal distribution of the data. This is a consequence of outliers, of which there are many in this case.

Although the graphical tools below doesn't confirm a normal distribution, there are reasons to believe that PU follows a normal distribution. This is based on the fact that plots in figure 1 does show signs of a normal distribution, combined with the existence of unusually high standard deviations of the PU ratings of some respondents. Therefore, it is assumed that the data of perceived usability are normally distributed.

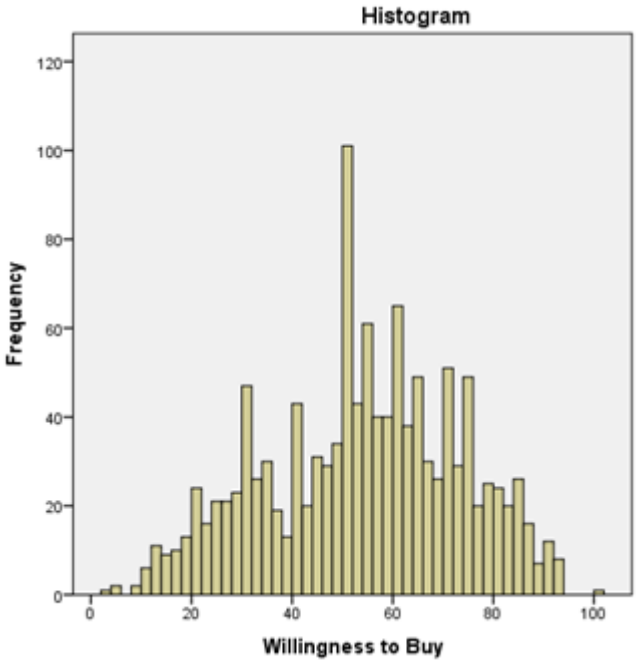
1.2 Willingness to Buy

In order to test whether the other dependent variable is normally distributed, the same procedure of tests for normality is carried out as in 1.1. In the case of Willingness to Buy, the Shapiro-Wilk test once again gave a test statistic which was significant at a 1%-level, suggesting a non-normal distribution. However, it was already concluded that this test was not suitable for large sample sizes. Therefore, graphical tools are once again used to identify the existence of a normal distribution. Figure 2 on the next page provides these different plots, along with some key statistics of willingness to buy.

Looking at the histogram of WtB, it can be noted that, in contrary to the histogram of PU, the distribution of the data seems to be considerably more symmetrical. This is confirmed by the skewness statistic (-0,194), which is over 4 times smaller than the skewness statistic of PU. This suggests that the data of WtB are less skewed and probably better fit a normal distribution. Another thing that catches the attention in the histogram is the high frequency peak for the ratings between 50 and 52. The simple explanation for this is that the rating of 50 lies exactly within this range (the actual range is 50-51,99). Since a rating of 50 is considered completely neutral in the used 0-100 scale, it is not surprising that this rating has been frequently chosen. However, since a rating of 50 wasn't as frequently chosen for perceived usability, this finding raises an important question. Were the respondents indeed completely neutral when rating the screens on WtB or did they have a lack of understanding of the question, leading them to rate the screen at a default, safe rate of 50? Unfortunately this kind of conclusions can't be derived from these findings. However, it is interesting to take this into account when developing surveys in future studies.

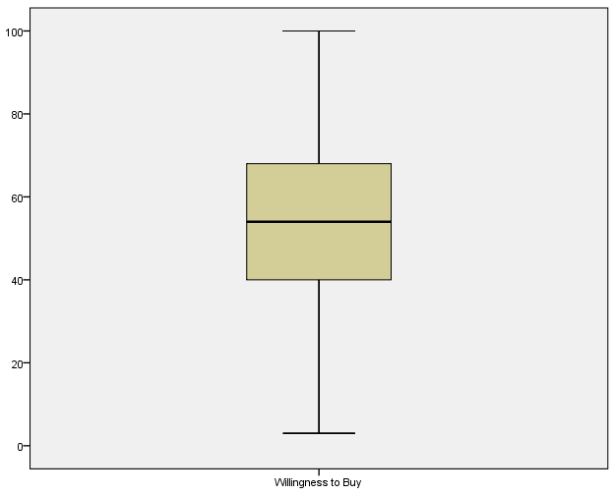
Nevertheless it seems like WtB is more normally distributed than PU. To further test this suspicion, the Boxplot and the Normal Q-Q Plot of WtB are investigated. Looking at the boxplot, it can immediately be noted that there are no outliers. Moreover, the quartile ranges of quartile 1 and 4 and quartile 2 and 3 are almost completely equal. This combination is typical for a normal distribution. The last graphical tool to test for normality of data is the Q-Q Plot. Comparing the Q-Q Plot of WtB with the one of PU, it can be noted that the points in the WtB Q-Q Plot tend to follow the straight line in the middle much better than those of the PU Q-Q Plot. Combining the findings of the graphical tests, it can be concluded that willingness to buy is normally distributed.

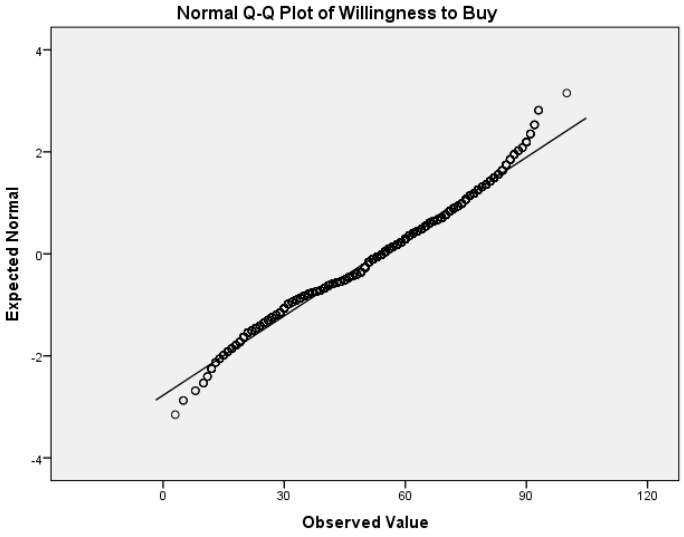
Figure 2: Graphical testing for normal distribution of WtB



		Statistic	St. Dev.
Mean		53,45	,549
95% Confidence Interval for Mean	Lower Bound	52,38	
	Upper Bound	54,53	
Median		54,00	
Variance		371,881	
Std. Deviation		19,284	
Minimum		3	
Maximum		100	
Range		97	
Interquartile Range		28	
Skewness		-,194	,070
Kurtosis		-,628	,139
Shapiro-Wilk Test		,986*	

1) * = statistically significant at 1%-level





2 Testing of Principal Assumptions of Multiple Regressions

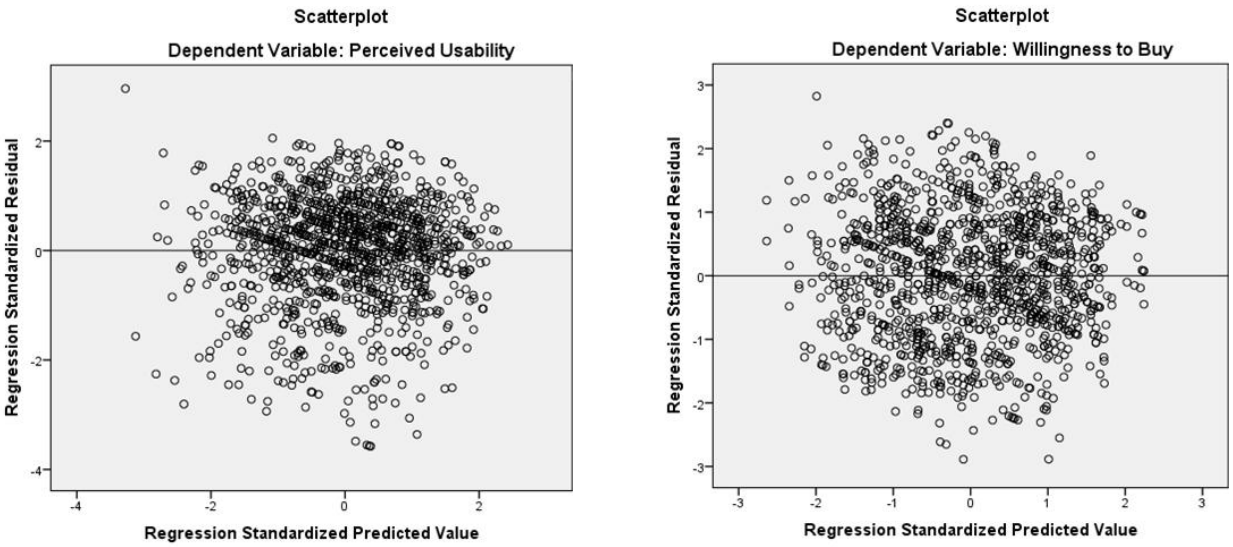
In general, there are four assumptions that must be met in order to justify the use of a multiple regression (Keith, 2014; Osborne & Waters, 2002), being:

- Independence of Errors
- Homoscedasticity
- No or little multicollinearity
- Normal Distribution of Errors

Each of these four assumptions will be tested for both dependent variables, perceived usability and willingness to buy, in order to justify the use of a multiple regression. Before the assumptions are tested, a test is performed to find the presence of clear outliers. In order to find the presence of outliers, the standardized residuals are plotted as a function of the standardized predicted values. What follows are the scatterplots shown in figure 3 on the next page, which are scatterplots of both dependent variables. In this study, it is assumed that measurements outside the [-3, 3] range can be defined as clear outliers that can possibly distort the data. Taken this into account, it can be concluded that there are no clear outliers in case of the dependent variable willingness to buy (see second scatterplot on next page).

However, the residuals in the first scatterplot, which represent the variable PU, are considerably different scattered. In part 1 of this appendix, it was already concluded that perceived usability contains numerous outliers below the average mean. The first scatterplot confirms the presence of a considerable number of outliers, most of them below the 0-line. These outliers also seem to lie outside of the acceptable [-3, 3] range. Although it is not a necessary procedure, it could be useful to track and delete the extreme outliers (those lying outside the acceptable range) in order to increase the validity of the model. In the table of figure 1 of Appendix C, a summary of the extreme outliers can be found. Eight outliers have been identified and subsequently deleted, resulting in a new scatterplot shown next to the table in figure 1 of Appendix C. This scatterplot already shows some signs of improvement when compared to the scatterplot below. However, there are still some clear outliers. Therefore, the same procedure is repeated, resulting in the deletion of five more outliers and the corresponding new scatterplot (Appendix C figure 2). After these adjustments, it can be concluded that the presence of clear outliers has been decreased. The dataset that remained after the deletion of the 13 outliers will be used from this point on for the dependent variable perceived usability.

Figure 3: ZRESID ZPRED Scatterplots



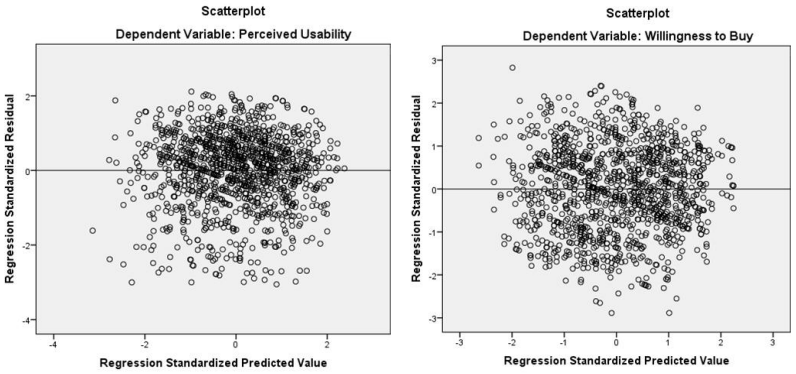
2.1 Independence of Errors

The first assumption that is made is that the errors/residuals are independent of each other, implying that the subjects are responding independently. If this assumption is violated, it can be concluded that the residuals are autocorrelated. In order to test this assumption, two tests will be used. The first test that is conducted is a statistical test, the so-called Durbin-Watson Test. This test is named after scientist James Durbin and Geoffrey Watson, who first presented this test in 1950 and 1951. The statistical value given by this test has a range of [0, 4], in which a value of 2 indicates no autocorrelation. When the value is lower than 1 or higher than 3, there are reasons to believe that the residuals are autocorrelated. The values for the Durbin-Watson Test are shown in the table in figure 4. For both dependent variables, the value is higher than 1 and lower than 3, which is deemed as acceptable. It must be noted though that the value of WtB is quite close to the limit of 1, which might raise some suspicion of autocorrelation.

However, the Durbin-Watson test is typically used in time-series data, for which it can be a powerful tool to detect autocorrelation. The data in this thesis doesn't include time-series though and therefore the D-W test might be less powerful to detect autocorrelation. Another way to test for autocorrelation is via a graphical analysis of the scatterplot in which the standardized residuals are plotted as a function of the standardized predicted values. As a rule of thumb, if a rectangular shape can be drawn around the residuals in the scatterplot, the assumption of independence of errors is met. For both PU and WtB, the shape that can be drawn around the residuals is roughly rectangular. This finding, in combination with the fact that the Durbin-Watson statistics weren't directly alarming, leads to the conclusion that the second assumption is met for both dependent variables.

Figure 4: Tests for Autocorrelation

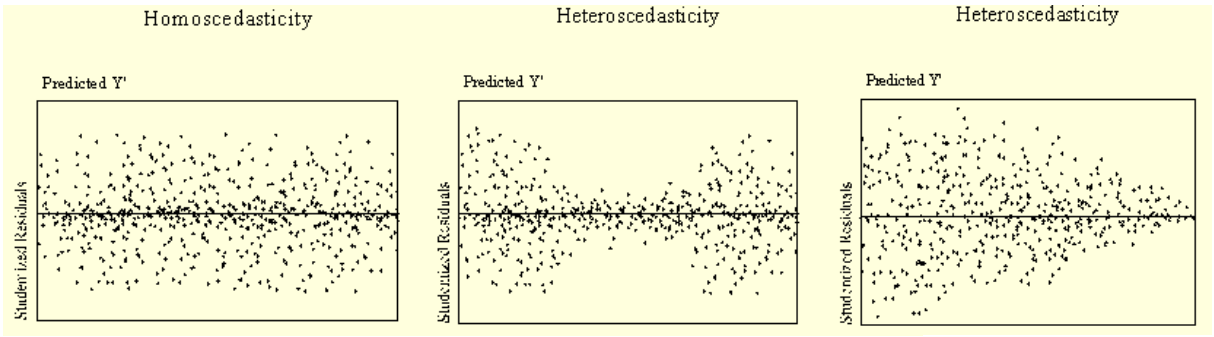
Durbin-Watson	
Variable	Statistic
Perceived Usability	1,445
Willingness to Buy	1,056



2.2 Homoscedasticity

The second assumption that has to be met, homoscedasticity, assumes that the variance of errors/residuals is equal across all levels of the independent variables (Osborne & Waters, 2002). If the variance of the errors is not equal across all these levels, it is marked as heteroscedasticity, which can lead to distortion of the findings and the statistical power of the analysis. In order to detect heteroscedasticity, the same scatterplots as used in the previous assumption can be examined. Ideally, the residuals in the plots are randomly scattered around the horizontal zero line, providing a relatively even distribution. Important is that the residuals do not follow a specific pattern. In case that a distinct pattern can be identified, it can be concluded that the variance of the errors is not equal across all levels of independent variables. In the figure below, homoscedasticity is compared with heteroscedasticity (Osborne & Waters, 2002). When the scatterplots of PU and WtB (figure 4, previous page) are visually examined, it can be concluded that no distinct pattern can be detected. Therefore, the assumption of homoscedasticity is met for both dependent variables.

Figure 5: Homoscedasticity explained



Osborne & Waters, 2002

2.3 No or little multicollinearity

The third assumption that must be met in order to justify the use of a multiple regression assumes that there is no or little multicollinearity. Basically, this means that the independent variables are uncorrelated. In SPSS, a widely used method to test for multicollinearity is to test a combination of two specific statistical measures: variance inflation factors (VFI) and Tolerance levels (Keith, 2014). In case that tolerance values are low in combination with large values for VFI (>10 is considered to be large), there is evidence of multicollinearity. Table 4 in Appendix C provides tolerance values and VIF values for each independent variables for both dependent variables. Based on these values, it can be concluded that there is no evidence for the presence of mulitcollinearity as there are no VFI values higher than 4. Therefore, the conclusion is that the independent variables are uncorrelated, which implies that the assumption of no multicollinearity is met for both dependent variables.

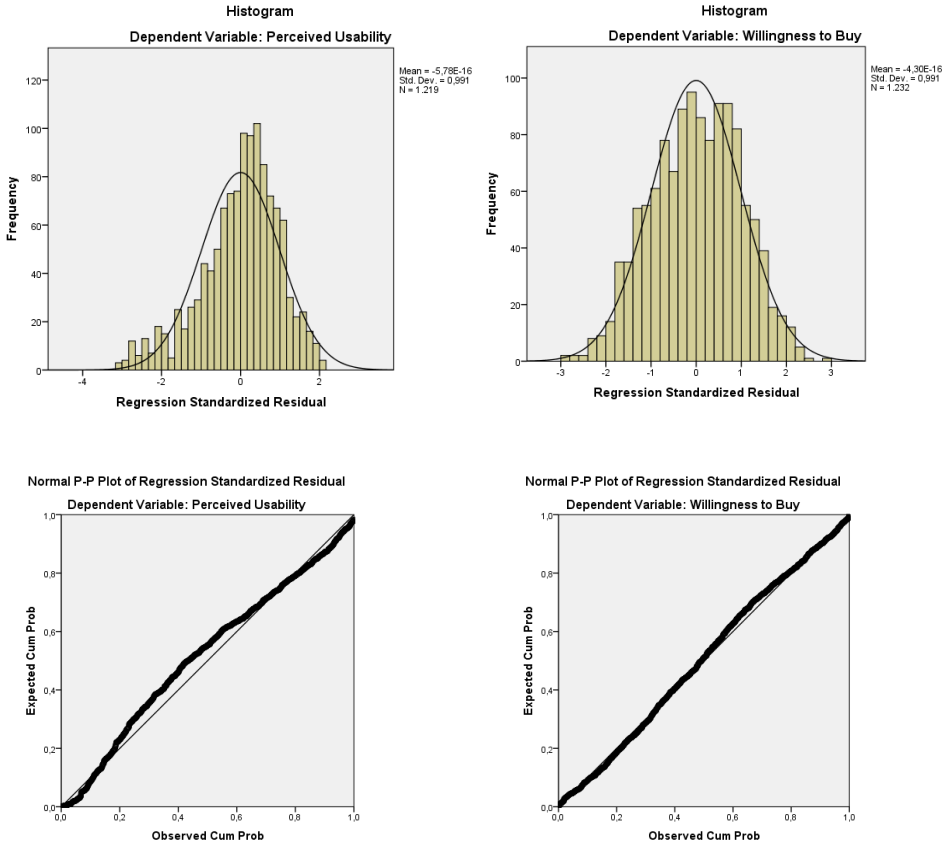
2.4 Normal Distribution of Errors

The last assumption of multiple regressions assumes that the errors/residuals are normally distributed. If this assumption is not met, relationships between variables and significance tests can be distorted (Osborne & Waters, 2002). There are several methods to test the normality of the residual distribution. In this case, histograms and Normal P-P Plots of Regression Standardized Residuals are visually examined to test for the presence of a normal distribution. Although the use of histograms is somewhat disputable in normality testing, it can still provide a useful indication of the presence of a normal distribution (Oztuna et al., 2006).

Figure 6 provides information about the histogram and P-P Plot for both dependent variables. Looking at the histogram and P-P Plot of perceived usability (left side of figure 6), it can be noted that the residuals in the histogram are somewhat negatively skewed and there is some positive kurtosis. However, the residuals do roughly follow the normal distribution line. The P-P Plot, which is considered to have more statistical power, shows that the data follows the straight 45-degree line, although it has to be noted that there are some deviations from this line. However, the deviations are not very large and therefore it can be concluded that the residuals of PU are normally distributed. When examining the right two graphs, it can immediately be noted that the variable WtB is clearly normally distributed, as the residuals follow almost exactly the normal distribution line (histogram) and the deviations from the 45-degree in the P-P Plot are very small.

This means that all the four assumptions of a multiple regression are met for both dependent variables. This justifies the use of multiple regression analysis, which is performed in the sections, 4.2.1 and 4.2.2.

Figure 6: Normality tests PU and WtB



Appendix F: Survey

Visual Design Effects

Q1 Thanks in advance for participating in this survey. This survey will consist of two parts. In the first part, you will be asked several questions about your experiences with ordering products online. In the second part, you will be asked to give your opinion about 16 different designs of an application. In total, this survey will take up about 15 a 20 minutes.

In case you want to have a chance on winning a voucher worth 25 euro, then please submit your mail address at the end of this survey.

(Part 1)

Q2 Gender

- ❍ Male
- ❍ Female

Q37 How often do you make online purchases via your smartphone?

- ❍ Never
- ❍ Seldomly
- ❍ Now and then
- ❍ Frequently
- ❍ Often

Q38 Indicate the importance of the following aspects when you purchase a product online

_____ Price
_____ Reviews of other users
_____ The ease with which I can find particular products
_____ Reliability of the website/applicatie
_____ The size of the product range
_____ The amount of product information

Q39 Do you currently own a food processor?

- ❍ Yes
- ❍ No

Q40 Would you be interested in buying a food processor?

- ❍ Yes
- ❍ No

(Part 2)

Q40 In the following part of the survey, you will be shown 16 screenshots of a smartphone. These are screenshots of an application of the online retailer Rakuten.com. With every screenshot you will be asked to give your opinion about the following two aspects:

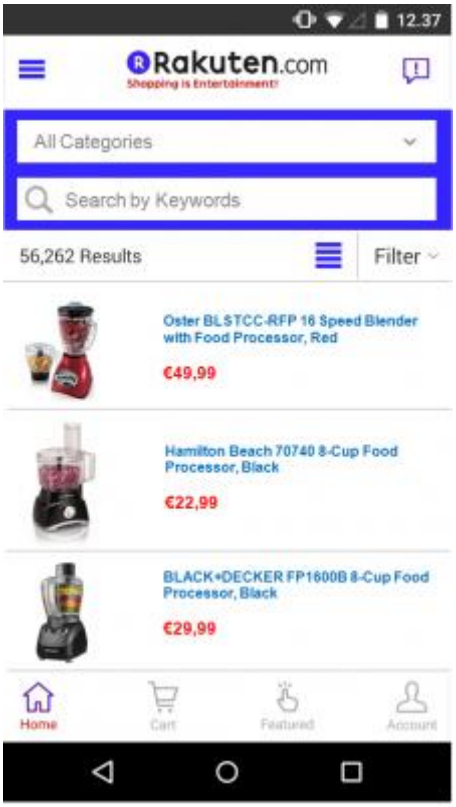
1. How user friendly/usable do you perceive this application: Imagine that you have to use this application. Would you experience the design of this application as user friendly and easy to use?

2. Suppose that you want to buy a food processor. To what extend would the design of this application motivate you to buy a food processor via this retailer?

It's important that you take your time and carefully look at the different screens before you proceed to answering the questions. When giving your opinion about the screens, you are expected to answer on a scale of 0-100, with 0 implying totally disagree and 100 implying totally agree.

The next questions (Q41 and 42(1) and 42(2)) are repeated 16 times, every time with another screen

Q41 Carefully look at the following screenshot and answer the corresponding questions



Q42 Indicate, on a scale from 0-100, to what extend you agree with the following statements (0 = totally disagree, 100 = totally agree):

_____ This design looks user friendly to me (1)
_____ I am motivated to buy a food processor by looking at this screen (2)