MASTER

Market research of a cloud-based quality analysis for mobile apps
a study at Omnext B.V.

Loeffen, M.T.C.

Award date:
2016

Eindhoven, May 3, 2016

Market research of a cloud-based quality analysis for mobile apps

A study at Omnext B.V.

by M.T.C. Loeffen

BSc Industrial Engineering and Management
Eindhoven University of Technology
Student Identity Number: 0721467

In partial fulfilment of the requirements for the degree of

Master of Science
in Innovation Management

Supervisors:

Dr. ir. W. (Michel) van der Borgh,	TU/e, ITEM
Dr. N. (Néomie) Raassens,	TU/e, ITEM
B.W. (Bryan) de Vries, Msc,	Omnext B.V.

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

Key words: market segmentation, decision making, software quality

Management Summary

Introduction

This research investigates how important the quality of mobile apps is according to various stakeholders within the mobile app market by performing a market segmentation. Omnext B.V., a company that provides technology and expertise dedicated to raising the quality and reducing the cost of software development and maintenance, is planning to extending its software analysis expertise toward the mobile software market by launching an independent cloud-based quality analysis service, referred to as an IMQA from here on. It appeared that the quality of current mobile applications (apps) is mainly defined in terms of usability and functionality towards the end users (Atkinson, 2013; Gerlich et al., 2015; Kim, Yoon, & Han, 2014; Maghnati & Ling, 2013; Wang, Liao, & Yang, 2013). An IMQA will analyze the technical quality of mobile apps. The output should provide developers and/or owners insights into the source code of their mobile apps in terms of maintainability, complexity, defects, risks, et cetera. To analyze which app developers and/or owners would be interested in an IMQA service, the following research question was defined: *How can segments in the direct customer base of Omnext be described in terms of the preferences and characteristics of an independent cloud based quality analysis for mobile apps?*

On the contrary, end users will not directly notice a higher technical quality of apps because an IMQA will not have an effect on the functionality of an app. Due to the fact that end users currently have to rely primarily on user ratings as signals to identify risky apps (Chia et al., 2012), this research took the idea behind quality labels into account which allowed to present a visible indication of the technical quality of apps in app stores, to test if it may influence the decision making process of end users. To analyze if different perceptions about app quality between end users exists, the second research question was defined: *How can segments in the indirect customer base of Omnext be described in terms of the preferences and characteristics of an independent cloud based quality analysis for mobile apps?*

Literature review

The theoretical analysis revealed that the mobile app market consists of four types of actors – app developers, app owners, app end users, and platforms that connect the actors – that are in some way related to each other. It was assumed that each actor would value the quality of apps differently and that each group of actors might be segmented in smaller clusters. It appeared that developers and owners show a lot of similarities regarding the behavioral and psychographic characteristics, because an app owner might be a developer at the same time. According to VisionMobile (2014), developers can be divided into eight segments which can be organized into three larger groups based on the goals each developer has: Self improvement, Revenues from the app economy, and Growing a business. On the other hand, an owner can be divided into two of the three groups found: Revenues from the app economy, and Growing a Business.

Varies studies (Hamka et al., 2014; Head & Ziolkowski, 2012; Plaza et al., 2011; Castells et al., 2004) had investigated the segmentation of end users based on the demographics and behavioral characteristics. In those studies, the end users were divided based on the type of usage. It became clear that the decision making process of end users differs when looking at the demographic and behavioral characteristics. Therefore, factors like age, gender, education, amount of mobile devices in use, and more, were used in the segmentation process of this research. Due to the scope, the *platforms* were not analyzed in detail.

Eventually, numerous attributes – factors that might have an influence on the decision making process of the actors – were derived from the previous studies about the mobile app market, and from interviews that were held. With extra input of two subject matter experts the following attributes were selected to use in the analysis: (1) Regarding developers and owners: Number of projects, Number of scans, Level of reporting detail, Independent quality label, Type of platform, and Price of an IMQA; (2) Regarding end users: Quality label, User rating, Number of reviews, App ranking, and Price of an app.

Methodology

The choice-based conjoint analysis (Louviere, 1988) was used to estimate the consumers' preferences by ask them to choose between two sets of alternatives that are prespecified in terms of different attributes and its levels. These choice sets were designed by using the software SAS and the final questionnaire was distributed online and on paper. The latent class analysis, used to estimate the part-worth utilities, was conducted with the software Latent Gold 5.1.

Analysis and Results

Unfortunately, the analysis was only performed on the survey that was focused on end users, because the survey meant for developers and owners received too few responses to perform an efficient analysis. In the end, data of 142 individuals was used to perform a latent class analysis. Based on the aggregate model, it was shown that end users could be segmented into four different classes. Besides, all five attributes appeared to have a significant influence on the purchase decision of end users. However, when taking the covariates into account, it was seen that only 'gender' and 'education' were significant. The first and largest segment (Peer followers), were mainly influenced by the user rating and the majority were men with a Bachelor's degree. The second segment (Balanced) was relatively equal influenced by the user rating as well as the price and consisted of as much men and women with a High school or Bachelor's degree. The third segment (Price-sensitive) obviously focused on the price and predominated by women with a High school degree. The last segment (Quality affected), seemed to be a small niche that were most influenced by a quality label. Besides, this segment consisted only of highly educated people.

Conclusion and Discussion

Overall, it was concluded that the majority of the end users were most influenced by the user rating and the price of an app. Still, it was interesting that a small group of highly educated people valued the presence of an acknowledged quality label more than the user rating. The fact that the research sample consisted mainly of Dutch students might explain that the user rating and price were highly valued in the purchase decision.

In general, the user rating is a very visible indication of the functional quality because it is presented on the front page in app stores. However, it has been discussed that this rating is also influenced by the performance of a mobile app. If an app crashes a lot or has other functional errors, end users will most likely start to complain. It is recommended that developers and/or owners quickly fixing what is wrong with the app. They might find the defects quickly when the source code is not very large or complex, however a lot of developers and/or owners have multiple apps or apps become very large and complex these days. In that case, an IMQA would be very helpful as it analyses the source code relatively fast and detects errors, defects or other mistakes in a short time span. Besides, personal feedback of respondents indicated that an interest in app quality exists, but that it depends what type of app it is.

Eventually, further research can be improved by using a larger sample including more variety of respondents to increase the significance of the covariates and be able to conduct a robust latent class analysis. In addition, the questionnaire sent to developers and owners should be simplified or other attributes should be taken into account.

Preface

This report presents the Master thesis of my graduation project that is conducted in partial fulfillment of the requirements for the Master of Science degree in Innovation Management. It concludes my master program at the faculty of Industrial Engineering and Innovation Sciences at the University of Technology Eindhoven. This research project is commissioned by Omnext B.V., who gave me the opportunity to perform my Master thesis in the form of an internship. I would like to express my gratitude to all the people that helped me with the conduction of this project.

First, I would like to thank my mentor and first supervisor Michel van der Borgh for all the guidance and assistance during my research period and for the feedback on the intermediate reports. His knowledge in the field of research helped me with defining the right scope, especially at the beginning, and successfully conduct a research that was scientifically relevant as well as interesting for Omnext B.V. Many thanks for the support during the decisions I could not make at first and eventually trust me in making the right decision. I would also like to express my gratitude to my second supervisor Néomie Raassens for her extra feedback and her knowledge of conjoint analyses. Her critical look on the intermediate reports, helped to focus the research and to design efficient choice designs.

Secondly, I am very thankful for all the support I received from my company supervisor, Bryan de Vries. He always made some time available to plan helpful meetings and to perform interesting discussions. Despite the difficulties in finding enough respondents, his enthusiasm towards this research kept me motivating to continue distributing the surveys. Furthermore, I would like to thank several employees to make some time available for being interviewed and to thank the CEO of Omnext to provide me helpful information about a cloud-based quality analysis service. Furthermore, I would like to thank several app developers and employees of Flitsmeister to provide me useful information about app development and by devoting some time for being interviewed.

Finally, I am very grateful for all the support of my family and my girlfriend for believing in me and assist me during the moments of adversity. The wise words of them kept me motivated to complete this project.

Mike Loeffen
Eindhoven, May 2016

Table of Contents

Chapter I – Introduction

This chapter contains the introduction to the market research of a cloud-based quality check for mobile applications. This research will be conducted in cooperation with Omnext B.V., hereafter referred to as Omnext. In the first chapter, a brief description of the company is provided. Secondly, the problem statement of the research will be derived in the second section. In the third section the relevance of this research will be discussed and in the final section the structure of this thesis will be outlined.

1.1 Company background

Omnext is a company that provides technology and expertise dedicated to raising the quality and reducing the cost of software development and maintenance. The company is founded in 2000 and positions itself as an independent partner specializing in the analysis of what are generally complex software systems. The aim of such analysis is the efficient development and maintenance of these software systems.

Organizations nowadays are extremely dependent on their software systems. They also need to be able to respond quickly to new developments (Time to Market), and there is a great need for flexibility. The source code of the software system forms the basis from which companies can respond to such new developments. It is therefore essential to be aware of, for example, the volume and quality of this basis if one wishes to act quickly and decisively. Omnext provides services for the automatic documentation and re-documentation of applications as well as services for measuring the quality and volume of these applications. Omnext enables organizations to gain – and stay – in control of the development and management of their – possibly outsourced – software. Omnext uses its own Omnext® SaaS technology for this.

The aim of Omnext is to become the recognized authority on realizing software quality and cost reduction within the markets where the company is active. Their goal is to achieve and maintain a leadership role in the field of quality and costs for software systems (Omnext.com).

1.2 Motivation and Problem statement

Software has become an essential part of everyday life of most people and companies. Costs of software development and maintenance can be high and maintenance costs are seen to increase linearly with an increase in the complexity of software (Banker et al., 1993). Lehman et al. (1997) demonstrated that software systems continue to evolve over time. In this process, software becomes more complex and starts to deviate from its original design, thereby lowering the quality of software (Mens & Tourwé, 2004). Companies and/or developers can avoid this, by taking action on time to reduce complexity and reduce the costs of maintenance in the end. Omnext, as a pioneer in software analysis, acknowledges the importance of software analysis to improve or indicate the software quality. An indication of the importance of software quality checks is that maintenance may span for 20 years, whereas development may be 1 to 2 years (Schach, 1999).

Besides the extensive use of general software worldwide, it can be said that software analyses of mobile applications become more and more important nowadays. Since the first successful smartphone launched on the market, the number of mobile devices drastically increased. According to 'State of mobile marketing' reports of Meeker (2015) the Internet usage of mobile devices in 2015 became higher than on personal computers. As a result, there is also a huge increase in mobile applications (apps). Despite that this indicates that mobile usage becomes more and more important in people's lives, still few research is done on the quality of mobile apps. Most research focuses on app usage from a user's perspective (Atkinson, 2013; Gerlich et al., 2015; Kim, Yoon, & Han, 2014; Maghnati & Ling, 2013; Wang, Liao, & Yang, 2013). This also applies to the current quality checks of mobile apps for devices running on Apple's iOS or Microsoft Windows before approval to their App stores. Google's Android has no mandatory quality check of mobile apps (Butler, 2011).

On the contrary, software providers, such as developers, owners or companies, have different interests and perceptions about software quality than end users (Jiang, Klein, & Discenza, 2002). Users might value the software quality more in terms of functionality, whereas providers might value it on the level of maintenance. Quality labels might affect consumer behavior, because an indication of the overall quality of mobile applications is difficult to appoint at the moment. Users have to rely primarily on community ratings as the signals to identify risky apps (Chia et al., 2012). However, these community ratings mostly reflect opinions about perceived performance or functionality and does not indicate how high the risk of this particular app can be. Providers, on the other hand, want to maintain and improve their competitiveness and controlling their costs by controlling software-related risks (Basili et al., 2010). Software quality analyses can improve the quality of software and decrease costs for providers. This will lead to a gap of the quality perception of mobile apps between developers/owners and end users because it is not clear if end users value this higher quality. Besides the fact that the mobile app market consists of these various actors, it is not completely clear if subsets, or segments, exist within each group of actors and, if so, how they differ according to various criteria based on geographic, demographic, behavioral and psychographic factors. In addition to the research about the segmentation of the mobile market, it is interesting to take the use of quality labels into account. Users find quality labels important in general, especially in selecting qualitative good food (Hildenbrand, Kühl, & Piper, 2015). However, it has not been tested extensively in the case of software. Besides, it might be that owners tend to select developers on the basis of their quality if a quality label is taken into account.

Omnext reports that a quality check of mobile apps could show potential according to current market trends. This is argued by the fact that the usage of mobile apps is increasing and software becomes more complex over time (Meeker,2015; Mens & Tourwé, 2004). On top of that, it appears that each group of actors might consist of different segments with different interests in the quality of apps. This research will put the theory to the test and fill in the gap in the existing literature from two different perspectives. The main problems mentioned are:

From the perspective of mobile app software providers:
How can segments in the direct customer base of Omnext be described in terms of the preferences and characteristics of an independent cloud based quality analysis for mobile apps?

From the perspective of mobile app end users:
How can segments in the indirect customer base of Omnext be described in terms of the preferences and characteristics of an independent cloud based quality analysis for mobile apps?

1.3 Research Questions

1.3.1 Software provider's perspective

1) What are relevant characteristics with which customer segments can be described?
2) What are the most important product attributes for customers when adopting an independent cloud based quality analysis?
3) Do different segments exist in Omnext's direct customer base on the preferences of an independent cloud based quality analysis?
4) What importance does each of the identified segments attach to the attributes of an independent cloud based quality analysis?
5) With which customer characteristics can each customer segment be described?

1.3.2 End user's perspective

1) What are relevant characteristics with which customer segments can be described?
2) What are the most important product attributes for customers when adopting an app?
3) Do different segments exist in Omnext's indirect customer base on the preferences of an app?
4) What importance does each of the identified segments attach to the attributes of an app?
5) With which customer characteristics can each customer segment be described?

1.4 Relevance

1.4.1 Academic Relevance

A lot of literature has been written about the importance of maintenance and quality of software (Mens & Tourwé, 2004; Schach, 1999; Slaughter, Harter, & Krishnan, 1998), however literature on the importance of *mobile* software quality lags behind. Though, mobile software could be written with the same programming language as other software, it often has other purposes than regular (enterprise) software and it is often simplified because mobile devices should be used on various devices and cannot handle a large amount of data (Kolinsky, 2014). The quality of mobile apps is mainly based on the functionality and is focused on a high usability. Developers regularly seem to neglect the technical quality, like complexity, risks or security issues, because end users would often not notice higher technical qualities, or because the Windows and Apple app stores screen the apps mainly from a user perspective, or the lack of experience (Inukollu et al., 2014). Research on mobile app quality has been done by Inukollu et al. (2014), where the researchers made suggestions for improvements of mobile app quality. Besides, it is not studied if user's perception of quality changes when a mobile app went through a quality analysis. Assigning a quality label to an app could influence the download behavior of users regarding apps, but this is until now mainly researched in the case of food (Hildenbrand et al., 2015). The results of this research can contribute to the field of mobile app quality analyses by expanding the research horizon by taking the interests of developers and/or owners of apps into account and by looking at the purchase decision making process of end users.

1.4.2 Practical Relevance

Nowadays, quality control of mobile software – in particular apps – is in most cases based on functionality from a user's perspective. In 2014, 22 percent of the total available apps had a low quality, like bad performance, security issues, or crashes at certain times (Inukolla et al., 2014). Different factors, such as the lack of expertise in terms of app development, unknown user demands and expectations, and poor maintenance can lead to low quality apps. Besides, out of the largest mobile platforms – Android, iOS, and Windows – only iOS and Windows have standard quality protocols for approving apps in their app stores. However, these quality controls are for most part based on quality perceptions of end users (Apple App review guidelines, 2015; Windows Store Policies, 2015). Most providers and end users of mobile apps are still not aware of the importance of high quality apps. High quality apps can reduce maintenance costs and could affect end user purchase behavior. Therefore, results of this research will be highly relevant for providers of mobile apps. In addition, by taking different stakeholders of mobile quality analyses into account it becomes relevant for Omnext too, as the company is planning to launch a cloud-based quality check for mobile software.

1.5 Research structure

This thesis is divided into five chapters. The first chapter was an introduction of the company in which the research will be conducted. It outlined the problem statement and relevance of the research. Following up on this introduction of the research, an extensive literature study will be conducted in chapter II. The literature study will focus on the perception of different stakeholders, such as developers and users, about mobile apps. Several attributes within the conjoint analysis framework will be defined. Information of this literature study will be used to come up with a research proposal in chapter III. This chapter will outline the main and sub research question that will be answered at the end of this research. In this chapter the research methodology, design and a detailed planning will also be provided. Chapter IV will include the analysis and results of the research followed by the conclusion and discussion in chapter V.

Chapter II – Theoretical Framework

This chapter discusses the theory behind the segmentation of the mobile app market and elaborate on factors that influences the decision-making behavior of direct and indirect actors of an independent mobile quality analysis (from here on referred to as IMQA). These factors will be derived from literature and interviews. Due to the exploratory design of this study, no hypotheses will be drawn.

2.1 Market segmentation

In order to sell a product, especially a new product, it is vital to understand consumer behavior and demand to reduce the risk of product failure. According to Cooper (1982), to increase the success of a new product, marketing resources are most critical. Market segmentation is generally accepted as being one of the key elements of modern marketing (Danneels, 1996; Kalafatis & Cheston, 1997; Wind, 1978). It is often used in multiple industries as a process by dividing the market into relatively similar segments of groups (Zikmund, 1999). Segmentation of the market is useful for two major reasons. First, it assists in analyzing the needs of a specific customer segment. Second, marketing campaigns can be focused more efficient and effective on these identified needs. For companies it allows to spend marketing and sales effort more wisely while meeting customers needs at the same time (Thach & Olsen, 2006).

2.1.1 Segmentation of different actors in the mobile app market (RQ1)

To segment the mobile app market, multiple levels and types of segmentation can be used. As there does not seem to be a scientific approach about segmenting the market into the most appropriate segments, it can be a challenging process (Oestreicher, 2009). One approach that is used more frequently and can be used for the mobile app market, is the needs-based market segmentation approach proposed by Best (2000). The basic criteria to segment a market are customer needs. These needs can by investigated by performing a market research. The different needs will be linked to consumer characteristics (Doyle, 2002).

The well-known classic marketing segmentation including four major categories of consumer characteristics will be grouped into:

- geographic; divides the market on the basis of geography, such as city or country;
- demographic; divides the market according to age, gender, income, education, etc.;

- behavioral; divides the market on the basis of specific behavioral patterns customers display when making a purchase decision, such as usage rate, benefits, spending, etc.;
- psychographic; divides the market based upon consumer personality traits, values, lifestyles, etc. (Crawford & Di Benedetto, 2010; Doyle, 2002).

The mobile application market consists of different types of actors. In the case of an IMQA it is assumed that these actors, either direct or indirect, will have different interests. During interviews with two subject matter experts of the mobile application market it is concluded that developers, owners, and end users are the most relevant actors for this research. It is important to note that regarding the decision making process of developers and owners in adopting an IMQA multiple people could be involved. Therefore, in segmenting these groups different roles should be assigned based on the *Decision Making Unit*. Relevant roles for this research will be an *influencer* and/or *decider* because these roles have a direct effect on the decision making process: Deciders make the actual purchase decision and influencers evaluates and recommends which purchase decision to make. It is assumed that developers will often be influencers as they have most knowledge about technical aspects of software and owners will often be the deciders.

Figure 2.1 shows the relationships that the different actors can have within the market as we know it today. Figure 2.2 shows a model which adds an IMQA to the market situation and illustrates how relationships will look like in such a scenario. The definitions of these actors are probably quite obvious. 'Developers' consists of people who develop a mobile application. Notice that this group can also be the 'owners' in case they deliver their mobile apps directly to the end users or app stores. Often developers are developing mobile apps for or within a company. Therefore, the company will be seen as the owner of the mobile app in this case and it explains the direct arrow between 'Developers' and 'Owners'. The arrow is pointed in both ways because a developer can be looking for or work for an owner, and an owner can have developers employed or looking for new ones. However, developers and owners can also be connected indirectly by an online platform. Such a platform can be seen as a marketplace to get developers and owners in contact with each other. Take for example an owner who has an idea of a new mobile app but does not have the resources to develop it. By using the online platform this owner can post a request for a developer to develop the new app. Developers, on the other hand, can search for requests and contact an owner in case the developer is capable of developing the app.

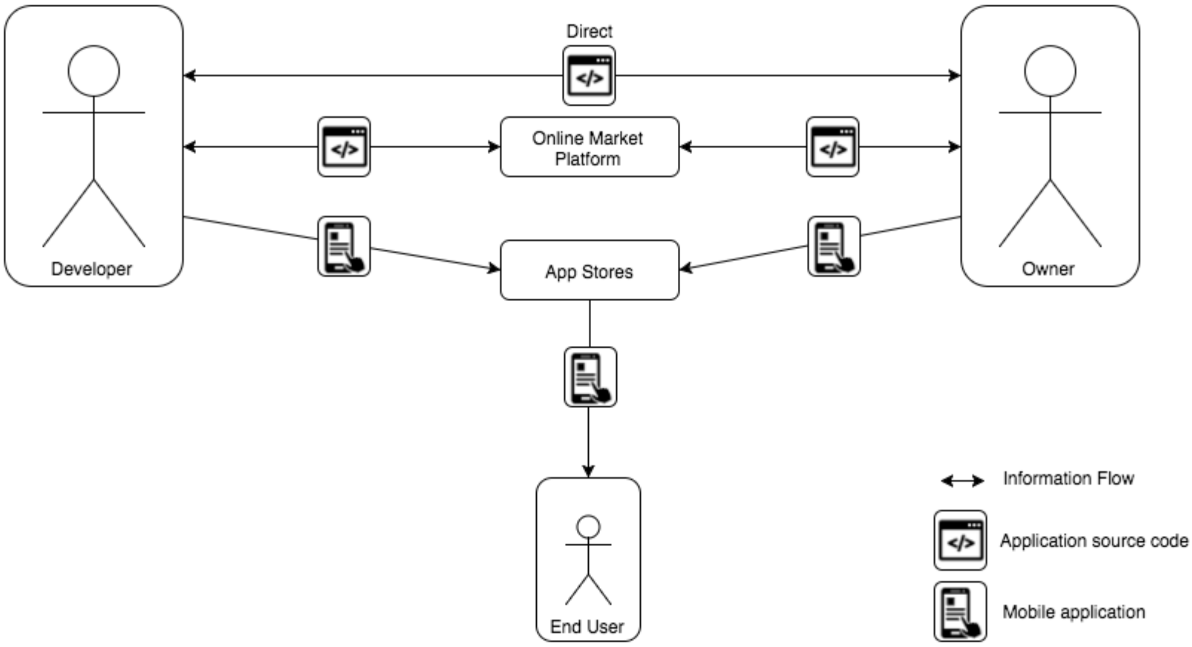

Figure 2.1: Relationship between actors

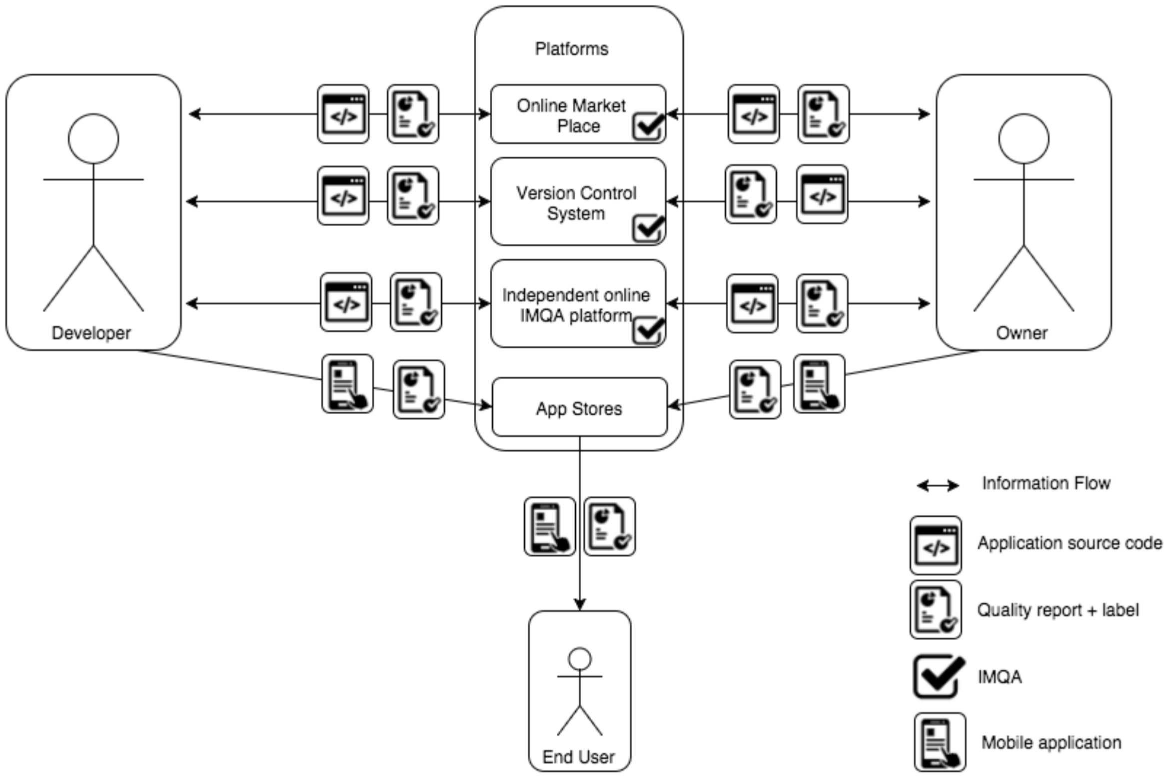

Figure 2.2: Market situation including an IMQA

Eventually, if a mobile app is ready for market the owners – or developers – submit it to a certain app store, like for example Apple App Store, Google Play Store or Windows store, where the end users can download the app for consumption. It is not very common that owners or developers choose to distribute their mobile app directly to customers, but it is possible.

In a situation where an IMQA is part of the mobile market, the model slightly changes (Figure 2.2). In one case, developers or owners can send their source code directly online to the cloud of an IMQA platform and eventually receive a quality report and quality label. In another scenario, they can merge an IMQA with a version control system. Version control is the management of changes to computer programs, large websites, documents, and other collections of information. In any development environment version management is considered as the most critical component and is essential to software development (Rapid Subversion Adoption, 2007). In combination with an IMQA the maintainability and development of the software could become better. In a third scenario, an IMQA could be an additional service of online market places. Because an online market place offers a platform to exchange source code between developers and owners, it can be interesting that during this process an IMQA is performed to indicate the quality of the software and the certain developer in general. After all these scenarios the developers or owners can submit their app, including a quality report and quality label, to the app stores. Therefore, end users could be made aware of the quality of the mobile application.

In general, as Figure 2.1 a simple overview of the division of the mobile market presents, the actors still represent very large groups, with diverse interests. Each actor probably has other goals. Besides, it can be expected that within these three groups the interests of each person are very diverge too.

2.1.1.1 Developers

Developers are important key actors on the mobile market, because they are the builders of the mobile apps and are therefore responsible for the failure of it (Inukollu, 2014). That is why the developers will be considered as one of the direct customers of an IMQA in this research. Often an app is engendered for a company's brand, product or service and is being used for marketing. According to its performance, it might have a great positive effect on brand equity or brand adhesion, but it could also be very destructive. For instance, when developing critical apps like a banking app customers have to be convinced regarding the security of such apps, but that can be very arduous. An IMQA might aid developers in delivering a higher quality app, and might increase confidence of customers in adopting an app. According to *The Developer Segmentation* report of VisionMobile (2014), developers of mobile apps can be divided into eight segments (Figure 2.3). This survey used data of more than 10,000 developers across more than 130 countries, so the geographic segment has no significant influence on developers. Additionally, due to the Internet, mobile apps can be easily distributed worldwide, thus physical borders are mainly diminished. Furthermore, the demographic segment does not apply to developers too. In this research the respondents will still be

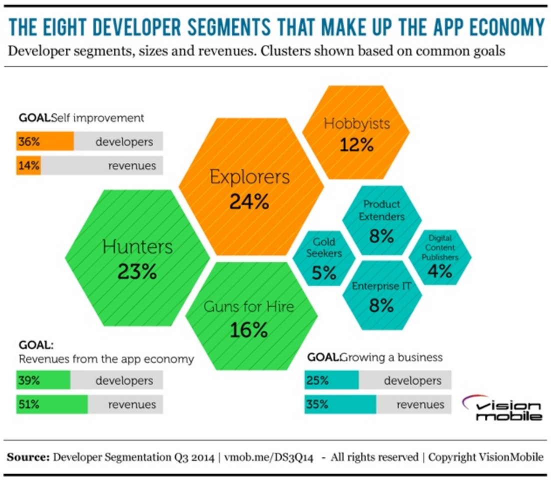

Figure 2.3: Developer Segmentation (VisionMobile, 2014)

asked about their geographic and demographic characteristics to analyze if there are differences in behavior regarding these aspects.

Most of the differences beforehand become clear in the behavioral and psychographic segment. Behavioral and psychographic features are different between developers as can be seen in Figure 2.3; two segments develop apps for self improvement, another four segments for growing a business and the last two seek revenues from the app economy. Despite these different goals, it is assumed that all types of developers could benefit from an IMQA as it can reduce development time and reduce (maintenance) costs. Therefore, it can provide feedback for self improvement, increase revenues or facilitate growing a business by increasing customer trust and brand value (Jones & Bonsignour, 2011).

2.1.1.2 Owners

Today, the developers of mobile apps are often not the actual owners of this app. Companies like King or Supercell, which developed multiple successful apps, consists of more than one developer. Besides, mobile apps like WhatsApp, that were invented by one person are often developed by external developers (Olson, 2014). Despite some differences compared to developers, owners of an app can be very similar segmented as developers. The main difference is that the two groups that aim for self improvement (Figure 2.3) are probably excluded. An owner – when not being a developer – needs developers to build the app. That is why the goal of self improvement cannot be applied to owners. An owner employs developers or it pays extern developers for developing apps. In the first scenario the owner can be considered as a company in most cases. Companies need to make strategic decisions for long-term benefits (Johnson & Scholes, 2002). Strategic decisions include competitive advantage, identify new possibilities for doing business, create new products, optimize revenues. Therefore, they can be segmented as direct customers of an IMQA in the groups aiming on increasing revenue or growing the business (Figure 2.3).

2.1.1.3 End Users

In the end, the end users are the group of actors for which the mobile apps are meant to use. These are the direct customers for developers and/or owners and therefore can be considered as indirect customers regarding an IMQA. According to a research about mobile customer segmentation of Hamka et al. (2014), the segmentation of this group is typically based on demographics and slightly on behavioral and psychographic segmentation. Geographic segmentation has no role in literature but will be analyzed in this research. Demographic segmentation indicates that younger end users are highly involved with their mobile phones (Walsh et al., 2010), while elderly people mainly use mobile applications to communicate with relatives (Plaza et al., 2011). In terms of gender, females mainly use their mobile phone as a fashion item and to maintain personal relationships, while men tend to use it for instrumental purposes (Castells et al., 2004). Psychographic segmentation found four segments consisting of unique needs, motivations, demands, and requirement on products or services or communication (Bouwman et al., 2008). From an application developer perspective, Hamka et al., (2014) found that end users can be segmented into six clusters. These clusters are presented in Table 2.1. It is remarkable that these clusters mainly consist of people older than 35 years. Another research analyzed only university students' behavior towards mobile applications with an average age of 20 years (Head & Ziolkowski, 2012). Their findings generated two distinct segments. Instant communicators, characterized by people who use their mobile phone for one primary objective: instantly communicate in a synchronous fashion. Other functionality, such as email, web browsing, music, and games were less used. Secondly, communicators/information seekers, consists of people who use the ability to communicate as well as information searching or gathering on the web. Head and Ziolkowski (2012), concluded that the instant communicators have a more hedonic perspective of their mobile phone, whereas communicators/information seekers perceive it as a utilitarian tool. In Table 2.1 these two segments are inserted. Important to note is that 'age' and 'education' characteristics do not count for these segments because they are in both segments equal, so only the amount of mobile usage is a relevant characteristic concerning the segments.

Table 2.1: Customer clusters from an application developer perspective

Clusters (Hamka et al., 2014	Characteristics	Segments (Head & Ziolkowski, 2012)
Application ignorant users	• Request a low number of URLs and make limited use of built-in and installed apps • Mostly females with age around 55-64 years old • Highly educated with full-time jobs and above average income level	Instant communicators • Average age: 20 • Highly educated
Basic application users	• Consult a limited number of URLs and make limited use of installed apps but run a medium number of apps • Mostly females with age around 45-54 years old • Highly educated with full-time jobs and above average income level	
Average app users	• Consult a medium number of URLs and moderate usage of downloaded and installed apps. • Mostly males with age around 35-44 years old • Medium to high educated with full-time jobs and above average income level	
Information seekers	• Request a high number of URLs but have a low usage of apps • Mostly males with age around 45-64 years old • Medium to high educated with full-time jobs and mostly above average income level	Communicators / Information seekers • Average age: 20 • Highly educated
App savvy users	• Request a high number of URLs and make extensive use of apps • Mostly females with age around 18-24 and 45-54 years old • Medium educated with full-time jobs and mostly above average income level	
High utility users	• Request a high number of URLs and make extensive use of downloaded apps but low usage of installed apps • Both males and females with age around 35-44 years old • Low to medium educated with full-time jobs and around average income level	

2.1.1.4 Platforms

Besides direct customers like developers and owners, an online platform that connects developers and owners with each other could be a direct customer too. If that is the case developers should upload the source code to the online platform so it can be retrieved by the owner. During this process an IMQA could be performed. Because such a platform is always aimed at connecting these two particular parties without being restricted to any borders, it is not necessary to segment these platforms because in general every online platform is similar to each other. In this research these online platforms will not be used as direct customers in the analysis. Furthermore, app stores could also make use of an IMQA. When focusing on the biggest players – Android, iOS and Windows – it is clear that they do not operate in similar ways. iOS and Windows have chosen to use a closed technology so they keep control over all strategic decisions including the approval of apps (Holzer & Ondrus, 2011). In contrary, Android is based on an open technology. Both technologies have advantages and disadvantages. Fact that the quality analyses of iOS and Windows mainly are focusing on functionality might indicate that a more elaborate IMQA is of interest. Especially in the case of Android which does not use any type of quality control. Though, these app stores might seem interesting customers for an IMQA, they probably are too big for the moment and will not be considered as direct customers in this research.

2.1.2 Decision-making process

As is discussed in the previous section, potential consumers of an IMQA can be divided into different market groups with different interests. Within each actor's group it became clear that segments exist. However, not many research has been done on the differences of interests between these actors. In general, the behavior of each actor will go through roughly the same process when facing a decision of adopting a new product. This decision-making process, consisting of five stages (Figure 2.4), was first introduced by Engel, Kollat and Blackwell in 1968 and is still a well used framework to evaluate consumers' decision-making process. It is important to mention that some of these stages may be skipped by consumers. Belch et al. (2012) adapted the model by taking the psychological aspect into account (Figure 2.5). In both models, differences between direct and indirect customers of an IMQA can be found.

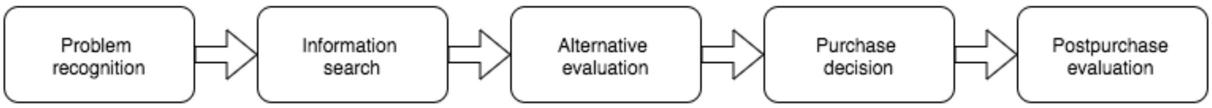

Figure 2.4: Stages in the consumer decision making model (Engel, Kollat and Blackwell, 1968)

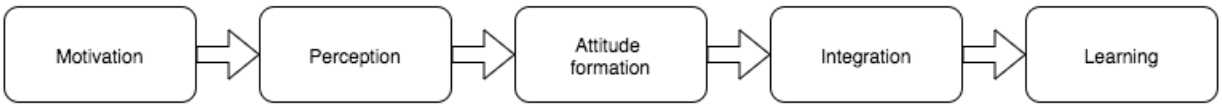

Figure 2.5: Relevant internal psychological processes in the decision making process (Belch et al., 2012)

According to the models, the initial stage starts when a consumer perceives a need and becomes motivated to solve the problem (Belch et al., 2012). It can be expected that the problem recognition between direct and indirect customers of an IMQA already differs. To understand these consumer needs and motivations Maslow's hierarchy of needs can be considered as a relevant approach to highlight five basic levels of human needs (Figure 2.6; Maslow, 1987). Based on the perspective of end users (Hamka et al., 2014), mobile apps can satisfy certain needs depending on the goal of this particular app including *Safety needs, Social needs, Esteem needs, and Self-actualisation needs.* It is important to note that *Safety needs* should be specified in software terms because Maslow (1987) specifies it in terms of personal or physical safety in general. In terms of mobile apps, in particular crucial apps like banking, it is very important that end users can use it safely and have the security that their privacy will not be harmed. A quality label might motivate end users reducing the boundaries of adopting certain apps or help them in selecting the best quality apps that belong in other levels of Maslow's hierarchy.

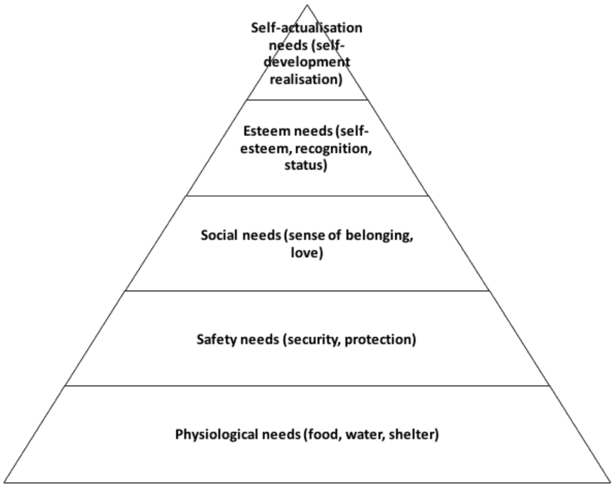

Figure 2.6: Maslow's Hierarchy of Needs

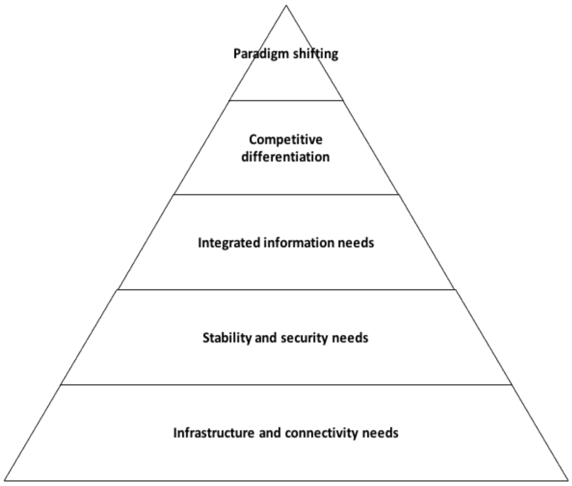

Figure 2.7: The IT Value Hierarchy

The direct customers of an IMQA – developers, owners or perhaps online platforms – will have other motivations and are difficult to assign to one of the five levels in Maslow's Hierarchy. Based on the perspective of Information Technology (IT), Urwiler and Frolick (2008) derived from Maslow's Hierarchy of Needs the IT Value Hierarchy (Figure 2.7). An IMQA satisfies different levels of the IT Value Hierarchy; *Infrastructure need* for adequate software and productivity, *Stability and security need* for stability and predictability in IT operations and transactional systems, and *competitive differentiation* to achieve a competitive advantage. As discussed in the previous section, an owner can be someone who has no knowledge of software and values more on the profit sector. This can result in unrealistic expectations of developers, like narrow schedules resulting in lower quality applications (Linberg, 1999).

In the second stage 'information search', differences between and within each actor of an IMQA can be expected. Nowadays, a lot of different information sources can be accessed and could influence the perception of each individual. Kotler et al. (2012) suggest that the most common information sources can be categorized in four different groups:

- Personal sources: family, friends, neighbors
- Commercial sources: advertising, sales people, websites, retailers, packaging

- Public sources: television, radio, newspapers, consumer organizations
- Experiential sources: handling, examining, using the product.

Due to the fact that almost everyone uses apps and most businesses have their own app, information about apps can be found everywhere. For developers and owners, information sources are limited because the range of different IMQA services are small and the target market is more specific. Therefore, most information is probably provided by sales people, websites or free trial services.

In this study, the main focus will be on the third stage of the decision making process. Before making the actual purchase decisions, consumers generally evaluate all relevant product alternatives based on the collected information. Again, it can be expected that this process differs between direct and indirect customers, as well as individuals within each group of actors. End users have a lot of reference possibilities in the app stores, whereas developers and owners have very few different IMQAs to compare. Moreover, this study tries to analyze which evaluation factors are most important to make the final purchase decision. In the end, a consumer decides to adopt the product that yields the highest individually perceived value based on the identified features, attitudes, and perceptions of the selected products (Belch et al., Kotler et al., 2012).

2.2 Attributes influencing the decision-making

The market segmentation indicates that between each actor different interests in mobile apps exist, especially between the direct and indirect consumers of an independent mobile quality analysis. What each actor values regarding a mobile application is already roughly presented but will be discussed in more detail in this section. The values each actor might have regarding an IMQA will be referred to as attributes from here on. For the purpose of this study, only a limited amount of attributes has been included. The attributes were selected based on (1) previous studies about the different segments existing in the mobile market, and (2) a qualitative pre-interview among some end users and mobile app developers. In Table 2.2 and 2.3 an overview is presented about potential attributes which will be discussed in more detail in the next two sections. These potential attributes are made *italic*.

Table 2.2: Potential attributes concerning developers and owners

Attributes	Attributes
Competitive Advantage	Number of Scans
Complexity	Overall Quality
Development Time	Performance
Functionality	Price of an IMQA
Independent Quality Label	Privacy Assurance
Level of Reporting Detail	Security
Maintainability	Size of Development Team
Number of Projects	Type of Platform

Table 2.3: Potential attributes concerning end users

Attributes	Attributes
Available Features	Performance
Control	Price of an App
Customization	Privacy
Feedback	Rating
Independent Quality Label	Reviews
Motivation to use an App	Security
Multi-Platforming	Usability
Novelty	Vividness
Number of Downloads	

2.2.1 Attributes derived from literature

The mobile market becomes more complex due to the growth and success of it. As a result, the mobile apps become more *complex* too, moving beyond inexpensive apps like simple games to more business-critical uses (Wasserman, 2010). Today, more than 1.5 million apps are available in the largest app stores focusing on different segments in this market. Therefore, each developer or owner might value a lot of different attributes depending on the function of a mobile app. As is concluded during the market research, the largest group of developers or owners focusses on optimizing revenues or business growth. To reduce the risk of failing apps developers and/or owners need to spend sufficient time on testing apps with respect to *security and performance criterion* (Inukollu et al., 2014). An IMQA might reduce these maintenance costs and development time to make it interesting for these stakeholders. However, due to the fact that it will be executed by an external company means the analysis comes with *a price.* But which *pricing levels or strategies* – like how many scans can be performed each year or how many projects can be analyzed? – would make an IMQA attractive to developers and/or owners? That is still unclear and might be relevant attributes in this research.

Besides decreasing costs, another goal of an IMQA is obviously *increasing the quality*. The quality of mobile apps can be referred to many different aspects, like *maintainability*, *security*, *privacy assurance*, *performance* or how *the app operates* as it should be. An interesting quote of Capers Jones (2011) – a specialist in software engineering methodologies and software qualities – about the trade-off between costs and quality is: "High-quality software is not expensive. High-quality software is faster and cheaper to build and maintain than low-quality software, from initial development all the way through total cost of ownership". Developers that can indicate that they deliver high quality apps, might also be more of interest to owners and might have a higher chance to be hired. On the other hand, an owner – or an independent developer – might want to increase quality to attract customers and build customer relationships, brand equity or brand adhesion (Inukollu et al., 2014). However, most end users will not notice bad software performance, bad security or privacy issues and value an app mainly on general performance. In the food industry it is proven that an *independent quality label* affects customer behavior positively (Hildenbrand et al., 2015), so it might affect end users of mobile apps as well. Moreover, developers could use these labels to indicate their quality in developing apps to become more of interest to owners. Eventually, increasing the quality and decreasing the development time, might result in a *competitive advantage* (Markovich & Moenius, 2013).

Moreover, crowdsourcing in the software industry, which is an emerging form of 'outsourcing' software development, is gaining significant attention (LaToza et al., 2013; Musson et al., 2013). Owners or developers could 'outsource' software development through different channels. As already is presented in Figure 2.1, owners could 'outsource' development directly or through an *online platform* like some kind of marketplace. 'Outsourcing' development has benefits, but also significant concerns (Stol & Fitzgerald, 2014), like task decomposition, challenging coordination and communication, and harder to maintain quality assurance. As suggested before, 'outsourcing' development directly to developers might be more attractive if an IMQA is used by these developers and might be an indication of the quality they deliver. However today, platforms provide a large online marketplace for developers and owners to meet (Stol & Fitzgerald, 2014). Most communication goes through these channels. In this case it is not yet researched if owners perceive more benefits if an IMQA would be used directly by developers or in collaboration with an online platform. Therefore, an attribute regarding developers and owners should analyze the preferred platform.

End users, on the other hand, value different attributes than developers and owners. Kim, Lin, and Sung (2013) find seven *design elements*, interpreted as related to engagement in their study: novelty, vividness, motivation, customization, feedback, control, and multi-platforming. These elements can also be interpreted as customer's perceived value. Monroe (1990) defines perceived value from a uni-dimensional perspective indicating that perceived value is a cognitive trade-off between benefits and sacrifices, like *quality* and *price*. This trade-off might be relevant in the case of mobile apps. Another approach, the multi-dimensional construct, includes more than two dimensions (such as perceived price, quality, benefits, and sacrifice) (Wang, Liao, & Yang, 2013). Mobile apps that underwent an IMQA should have a higher quality, or lower risks, that might increase the attractiveness of paying for or adopting apps.

However, as an IMQA focuses on the source code of a mobile app, most end users would not notice the effects of a good source code because an end user will generally experience if an app does what it has to do. A good source code is in most cases mainly relevant to its maintainers, because it has a better readability to improve maintenance and increase the efficiency of development. Regarding end users, that could mean the app has a better *performance* overall or during launch, and updates could be released quicker to maintain a good performing app. However, as stated before, using a *quality label* on apps that have a good source code with the help of an IMQA might be easier noticeable by end users and affect their *perception about the quality* of the app (Hildenbrand et al., 2015). In general, most end users download an app based on other users' *reviews* in terms of a *rating* or *number of downloads.*

Finally, *security or privacy risks* can be a factor of quality from the perspective of end users. Due to the increasing number of APIs available gives rise to new types of these risks (Sadeh & Hong, 2014). Malware becomes more and more of a problem (Burguera et al., 2011; Felt et al., 2011). Besides, end users are often unaware of the amount of information the apps access and why (Sadeh & Hong, 2014). If an IMQA might decrease these risks, it could increase the quality (perception) for end users.

2.2.2 Attributes derived from interviews

To become aware of some actual data from the market, a couple of small interviews were performed among developers and end users. The questions and answers are presented in Appendix I. From these interviews it can be concluded that some factors match with the information derived from literature. For example, the developers and manager mentioned that they base their app *quality* on end users' demand and feedback. But, they all agreed on the fact that *development* is the main expense of mobile apps.

Besides, it becomes clear that the developers and manager see the potential of an IMQA, but they stated that it probably will depend on the *goal or function of a mobile app* and the *size of the team developing* the app. In case of large enterprises that offer risk full products, like assurance companies or banks, an IMQA could be helpful to reduce risk. In addition, companies which consists of a large team of developers, an IMQA could be helpful to increase the maintainability of the software. Also the fact if a developer is intern or extern might play a certain role. The developers of the interviews could imagine that a mobile quality analysis has benefits in the case a company outsources some businesses. A quality label might be a helpful indicator in selecting a good developer. Eventually, the developers and manager are willing to adopt an IMQA if it could be proved that it actually has benefits for their business and improves the maintainability of their software. This includes the process of the *final reporting* in terms of how useful the reporting is in helping to improve the software of a mobile app.

Regarding end users, the developers and manager were skeptical if the end users could notice the benefits of an IMQA as well. In the end, the *performance* of a mobile app should be better, however in terms of *usability* and *available features* nothing will significantly change. Therefore, a *quality label* might be used to make end users aware that certain mobile apps can be trusted in terms of low risks for example. According to one developer this only might have success if the quality label has a high positive reputation.

Moreover, in the interviews the end users confirmed that the quality of an app will be relevant when considering to adopt an app. Though quality is a very broad concept, one end user explained '*privacy*' is an important factor when considering the quality of an app. According to two end users the quality – if referred to *privacy, security* and *performance* – will be a factor in making the decision to pay for a certain app. Besides, when asked if a quality label would affect their quality perception, they also indicated that it depends on the reputation of the label.

Eventually, due to the very small group of interviewees, the results are of course very general. It is imaginable that the actual market is much broader. As can be concluded from the interviews, the potential of an IMQA depends on the behavior and goals of the end user and developers or owners.

2.3 Summary

Three different actors have been identified that could have a direct or indirect interest in an independent mobile quality analysis (IMQA). Each of these group of actors has been segmented to analyze the different factors that have an influence on the decision making process of adopting an IMQA – in the case of developers and owners – or a mobile app – in the case of end users. Based on different theories about mobile app development and usage, and pre-interviews, different attributes have been identified that might have an influence on the decision making process of developers and owners, and another set of various attributes have been identified that might have an influence on the decision making process of end users.

Chapter III – Methodology

This chapter will describe and discuss the methods that have been used while conducting this research. First, the research design will be discussed. Afterwards, the conjoint analysis will be explained by elaborate on the various steps involved in this process.

3.1 Research method

As mentioned in the previous chapters, the basis of this research is exploratory by testing the relative importance of each attribute and identifying the most important selection criteria in the decision making process of direct and indirect customers. In order to identify these 'preferences' of each customer, different models exist. Green and Srinivasan (1990) consider three alternative approaches for measuring preference structures: compositional, decompositional, and hybrid. A compositional approach, like the Self-Explicated Approach, can be used if there are a large number of attributes. People will be asked to evaluate each attribute (level) independently. In contrast to compositional approaches, decompositional approaches, like conjoint analyses, works very well when there are only a few (six or fewer) attributes. In this case respondents react to a set of total profile descriptions (Sattler & Hensel-Börner, 2000). Finally, the hybrid approach combines the ease of administration of compositional models with the greater amount of realism of decompositional models (Green & Krieger, 1996). Due to the fact that in this research no more than six attributes will be analyzed, a decompositional model is preferred (Sattler & Hensel-Börner, 2000). Therefore, a conjoint analysis will be used.

3.2 Conjoint analysis

A conjoint analysis is a statistical technique to study the factors that influence consumers' purchasing decisions. According to the developers and pioneers of the analysis, Green and Srinivasan (1978), a conjoint analysis is defined as "*any decompositional method that estimates the structure of a consumer's preferences (i.e., estimates preference parameters such as part-worths, importance weights, ideal points), given his or her overall evaluations of a set of alternatives that are prespecified in terms of levels of different attributes*". In general, a respondent is asked to choose between two or more scenarios. However, including more choice sets, increases the choice complexity (DeShazo & Fermo, 2002). Each scenario consists of the same attributes but with different levels assigned. An attribute is a characteristic of a certain product or service, made up of

specific levels. By observing which scenario each respondent chooses, it is possible to estimate the utility each attribute level has regarding overall product preference (Louviere, 1988).
A conjoint analysis can be carried out in many ways. Green and Srinivasan (1990) provided an overview of alternative methods across the different steps involved in a conjoint study. However, this overview is not complete, since it does not take all alternative methods into account, like mixed preference models and mixed collection methods. It is also lacking the number of alternate constructions sets and estimation methods. Therefore, the overview of the different steps involved is slightly adapted in this research. The adapted framework of the different steps involved in this conjoint study is presented in Table 3.1.

Table 3.1: Steps involved in conjoint analysis

1.	Selection of the type of conjoint analysis
2.	Selection of the attributes
3.	Selection of the specific levels per attribute
4.	Creation of the survey
5.	Data collection method
6.	Estimation of part-worths
7.	Apply latent class analysis

3.3 Selection of the conjoint analysis and attributes (Step 1 – 3; RQ 2)

In this research is chosen to use the choice-based conjoint analysis because it best suits the purposes of this research. The respondent is asked to indicate the option or scenario they prefer in multiple alternatives instead of rating each alternative. This technique allows researching a customer's utility function by calculating part-worths, importance weights and ideal points and provide managers insights in customer choices (Hair et al., 2010). It is argued that respondents have less difficulty in choosing one product out of two rather than rate all the alternatives independently (DeSarbo et al., 1995). Therefore, it is seen as a more realistic and relatively simple task for respondents. Thus, the choice sets presented in a choice-based conjoint analysis are very similar to realistic choices faced in the marketplace.

In general, the selection of (the number) of attributes depends on the type of conjoint analysis. The preferred number of attributes when using the traditional Full-Profile or Choice-Based Conjoint Analysis should be six or fewer, while the Adaptive Conjoint Analysis is preferred for larger number of attributes (Hair et al., 2010; Orme & King, 1998). In Table 3.2 and 3.3 the selected attributes and their corresponding levels are presented. The requirement that the maximum number of attributes in a choice-based conjoint analysis should be six (Hair et al., 2010), is met in this research. Too many attributes lead to a situation where respondents will not react in a proper way, but using too few attributes is not recommended because the researcher might exclude some important characteristics. Additionally, it is suggested to use equal number of levels for all the attributes to exclude the Number of Levels effect (Virelli, 2001). This effect might occur when one attribute has more or fewer levels than other attributes resulting in a situation that an attribute with more levels appears to be more important. The attributes and levels were selected based on (1) previous studies about the different segments existing in the mobile market, (2) the input of two subject matter experts, and (3) a qualitative pre-interview among some end users and mobile app developers. In Appendix II the definitions of each attribute are provided.

Table 3.2: Selected attributes and levels concerning developers and owners

Attribute	Level 1	β1	Level 2	β2	Level 3	β3
Number of Projects (/year)	1 project	-1	5 projects	0	10 projects	1
Number of Scans (/year)	4 scans	-1	12 scans	0	More than 12 scans	1
Level of Reporting Detail	Low	-1	Medium	0	High	1
Independent Quality Label	No label provided	-0.66	Label provided but **not** acknowledged by market	0	Label provided and acknowledged by market	0.66
Type of Platform	Independent online analysis platform	-0.33	Version Control System	0	Online Market place	0.33
Price of an IMQA (/month)	120 dollar	-1	60 dollar	0	30 dollar	1

Table 3.3: Selected attributes and levels concerning end users

Attribute	Level 1	β1	Level 2	β2	Level 3	β3
User Rating	1 star	-1	3 stars	0	5 stars	1
Number of Reviews	10.000 reviews	-0.33	500.000 reviews	0	1 million reviews	0.33
App ranking based on number of downloads	Top 50	-0.33	Top 10	0	Top 3	0.33
Quality Label	No label provided	-0.66	Label provided but **not** acknowledged by market	0	Label provided and acknowledged by market	0.66
Price of a mobile app	4.99 dollar	-0.66	2.99 dollar	0	0.99 dollar	0.66

The prices of a mobile application are based on two statistical studies about average prices of apps in the Apple App Store (Pocket Gamer, 2012; Statista, 2015). It is assumed that the prices of similar apps in Google Play Store or Windows Store are equal to prices of apps in the Apple App Store. The prices of apps used in this survey are derived from these statistics and slightly adapted to more common displayed prices.

3.4 Design and data collection method (Step 4 – 5)

The data collection method that is used in this research is a survey including a questionnaire. Two different surveys are made because this research is intended to explore two different aspects regarding the quality of mobile apps. One survey is designed to explore the perception of mobile app developers and/or owners in regards to the different attributes of an independent mobile quality analysis. The other survey is designed to explore the perception of end users of mobile apps in regards to other various attributes that are likely considered for using an app. This means that this research consists of two different target populations. The target population regarding the survey intended for the developers and owners includes any mobile app developer or owner, where the latter could be one person or multiple persons in the case the owner is a company. Regarding the survey intended for end users of mobile apps, the target population includes any consumer of mobile apps. In both researches the respondents could have any nationality. Additionally, individuals of any gender and any age are taken into account. So, this study is defined as a cross-sectional study, because it analyzes the possible effects of various groups at a single point in time and sort out the effect of multiple independent variables (attributes) upon the dependent variable (decision-making process).

The final design of both surveys are presented in Appendix IV. The surveys are in English because the target group includes any app developer, owner or end user around the world. Both surveys are pre-tested on a sample of 5 participants each, meaning that 5 developers and/or owners of mobile apps were involved in the pre-test regarding the preferences of an IMQA and 5 end users of mobile apps were involved in the pre-test regarding the preferences of mobile apps. These participants were excluded from the actual data collection. Eventually, it is decided to place the attribute 'Independent Quality Label' regarding the questionnaire meant for end users at top in each scenario.

3.4.1 Structure survey

The structure of the two surveys are similar and start by introducing the purpose of the study. After the introduction, 18 different choice sets are presented and the respondent is asked to select which of the two alternatives they prefer. These choice sets are different in both surveys because other attributes are used. A total of 729 alternatives are possible in the case of the survey for developers/owners, and a total of 243 alternatives are possible in the case of end users. To make efficient choice designs the final selection of the most relevant attributes and levels should be characterized by four requirements (Huber & Zwerina, 1996):

- Level balance: the requirement that the levels of an attribute occur with equal frequency.
- Orthogonality: this condition is required when each level of an attribute occurs with each level of another attribute with equal frequencies.
- Minimal overlap: the requirement that the probability that an attribute level repeats itself in each choice set should be as small as possible
- Utility balance: the criterion that the probability of each alternative in the set being chosen are approximately equal.

The sets were designed by using the statistical software SAS. The approach of SAS results, in almost every case, in efficient designs based on the four requirements (Kuhfeld, 2010). The entire process of the design can be seen in Appendix III. Though, the true β-values, which indicate the level of influence on the decision making, are unknown, it can be assumed that the attributes are not equally weighted. The assumed β-values for each attribute are based on the input of two subject matter experts and interviews with certain mobile app developers and mobile app end users. The β-values are listed in Table 3.2 and 3.3. The higher the value, the higher the assumed influence the certain attribute has on the decision making process. Eventually in the last part of the surveys, the respondent is asked to answer socio-demographic questions regarding the covariates: age, gender, nationality, education, profession, and income.

3.4.2 No-choice option

In the survey intended for app end users a no-choice option is included. However, in the other survey intended for app developers and owners the no-choice option is *not* included. In literature, several reasons why respondents would choose the no-choice option can be found (Vermeulen, Goos & Vandebroek, 2008). The rationale theory states that a consumer prefers the product that offers him the maximum amount of utility. In reality it is possible that a consumer needs to choose between several alternatives but none of these alternatives is considered as attractive. So, in that case the consumer probably chooses the no-option and look for other alternatives. Another theory, provided by psychological research, indicates that consumers often try to avoid intricate trade-offs and could sense discomfort and fear of making the wrong choice (Baron & Ritov, 1994). In this situation the no-choice option is used as a way to avoid difficult decision making, which is possible when both alternatives are nearly equal or when both alternatives are unattractive. The decision to include a no-choice option in the survey intended for end users of apps, comes from the fact that apps nowadays are widely used. People know what apps are and how to use these, so most have already experienced a decision making within the app market. Therefore, to create a more realistic scenario and to assume most end users will perceive low difficulty in the decision making of mobile apps, it is decided to include the no-choice option in the survey intended for mobile app end users. However, the survey intended for developers and/or owners of apps will not have a no-choice option. The reason for this is that an independent mobile quality analysis is relatively new on the market which can lead to a situation where the respondents may not be interested in the product and continually choose the no-choice option (Haaijer, Kamakura, & Wedel, 2001). If that would be the case than it would be very difficult to analyze the results as it will not become clear why respondents have chosen the no-choice option. Because one goal of this exploratory research is to analyze which alternatives of an IMQA are most attractive, it is decided to exclude a no-choice option and force respondents to choose the 'best' alternative in the case of the survey intended for developers and owners of mobile apps.

To test if the assumptions made lead to a satisfactory outcome, the survey was sent out in the form of a pre-test. With the feedback, slight adaptions were made and the final survey was distributed on various online platforms and send via mail to random selected mobile app companies. Besides, to increase the response rate, the complete survey is developed according the guiding principles of Dillman (2009). Due to the broad respondent population, snowball sampling (Malhotra & Peterson, 2006) by e-mail is created by asking respondents to forward the survey invitation to others through online channels, such as e-mail or social media.

3.5 Statistical Design (Step 6 – 7)

To get answers to the research questions, the gathered data from the surveys will be analyzed with the conjoint analysis. The objective of the conjoint analysis can be achieved by estimating coefficients called part-worths (or utilities) for the various attribute levels by decomposing the measured overall preferences for product profiles into the part-worth utilities (Manuel, Adrian, & Benavent, n.d.) The conjoint analysis will be applied with the estimated part-worth utilities provided by a latent class analysis by using the software Latent Gold Choice 5.1.

Chapter IV – Analysis and Results

This chapter first describes the preparation of the data and the descriptive statistics. Secondly, the latent class analysis of the obtained data will be discussed by using the software Latent Gold Choice 5.1 and eventually with the estimated part-worths a conjoint analysis is applied. After segmenting the data, the consumer characteristics are combined with these segments to be able to answer the research questions.

4.1 Preparation of the Data

After the pre-versions of the surveys, only the text sections were adapted to make it less long and more readable. The actual choice sets and other questions remained the same. On February 3, 2016 both surveys were launched on the online platform SoGoSurvey. Due to a very low response rate it was decided to extent the data collection period and in addition to the online platform paper-and-pencil versions were used. Eventually, the data collection of both surveys was closed on March 3, 2016. A total of 153 participants have completed the survey that was meant for mobile app end users. Unfortunately, in the case of the survey that was meant for mobile app developers and owners, only 9 people filled out the questionnaire. Though the survey was published on many different channels and statistics showed that many people opened the survey, still too few people completed the survey. Due to the lack of time and the risk of falling behind schedule too much, it is decided to stop the survey after 4 weeks and exclude the received data of developers and/or owners from this research. So, from here on only the results of the end users will be used.

Of the 153 participants that filled out the survey, one entry is deleted on first sight. This individual answered 'no preference' for all eighteen choice sets and its demographics were also very unusual compared to all other responses. The remaining data set of 152 respondents did not contain any missing values. To detect observations that are distinctly different from the majority of the data set, an outlier analysis was performed. A variable with an unusual high or low value or a unique value are seen as outliers and may distort the results of this research (Hair et al., 2010). It is recommended to exclude the outliers. Univariate outliers, which are observations having high or low values within a single variable, were identified by converting the data of each variable into standard scores. According to Hair et al. (2010), the threshold value of standard scores for large sample sizes (80 observations or higher) is ±4. Based on this rule of thumb, ten observations were identified as outliers. Before deleting the data immediately, it is necessary to analyze the uniqueness compared to other observations. After this process, it is decided to delete these ten observations. Next, the data set was examined for multivariate outliers with the Mahalanobis D^2 measure (Hair et al., 2010). Five respondents were seen as outliers because they had a D^2 with a probability less than 0.001. This time, it is decided to keep the data of these respondents in the data set because each variable seemed valid enough compared to the remaining observations. In detail, the responses on the eighteen choice sets of these five outliers were acceptable as the ratio of the 'no preference' response was normal and the remaining 'A' or 'B' responses were not contradicted to the other observations. Besides, the demographics of these individuals showed no extreme values. The data set eventually contained 142 observations.

4.2 Sample

In Appendix V the complete descriptive statistics of the 142 respondents are presented. A total of 28.9% of the respondents were from foreign countries. These foreign respondents represented a very small sample in their own country, therefore they have been combined in larger regions. Other European and North-American respondents were slightly older than the average age. The South-American and Asian respondents were slightly younger. Regarding the number of devices used, it can be seen that Dutch respondents on average use 2 devices each week whereas foreign respondents have a mean of 1.50. Furthermore, almost all respondents have finished college or have a university degree. With respect to their current profession more than 50% were students and more than 25% had a full-time job. That the majority of the respondents were students probably explains that most respondents had a monthly income of €1500 or less. The usage of iOS compared with the income had approximately the same mean as Android compared with income. Only respondents who use both iOS and Android had a higher mean income.

4.3 Latent Gold Analysis

The latent class analyses were conducted using the software Latent Gold Choice 5.1. Before performing this analysis, it is required to check the fit of the β-values of the independent variables that were derived in Chapter III and that were used in the software SAS to create the questionnaire. In Table 4.1 the assumed β-values are compared with the actual choices of the respondents.

Table 4.1: Goodness-of-fit analysis of the derived Beta-values

Set	Choice	Expected probability by SAS	Actual percentage of choices by participants
1	A	72.71%	100%
	B	27.29%	0%
2	A	42.07%	0%
	B	57.93%	100%
3	A	49.75%	19.05%
	B	50.25%	80.95%
4	A	58.66%	96.92%
	B	41.34%	3.08%
5	A	49.75%	91.18%
	B	50.25%	8.82%
6	A	73.11%	30.30%
	B	26.89%	69.70%
7	A	34.30%	88.06%
	B	65.70%	11.94%
8	A	41.58%	93.98%
	B	58.42%	6.02%
9	A	49.75%	2.90%
	B	50.25%	97.10%
10	A	42.07%	66.07%
	B	57.93%	33.93%
11	A	41.58%	97.12%
	B	58.42%	2.88%
12	A	57.93%	13.85%
	B	42.07%	86.15%
13	A	41.58%	37.70%
	B	58.42%	62.30%
14	A	42.07%	53.73%
	B	57.93%	46.27%
15	A	79.25%	86.73%
	B	20.75%	13.27%
16	A	41.58%	26.49%
	B	58.42%	73.51%
17	A	79.25%	61.36%
	B	20.75%	38.64%
18	A	73.11%	87.22%
	B	26.89%	12.78%

In the goodness-of-fit analysis of the β-values in Table 4.1 can be seen that most sets vary greatly compared to the expected probability. However, a small majority – 10 out of 18 – of the sets still predicted the right directions of the answers of the participants. For example, SAS estimated that Choice A will be chosen more often than Choice B in Set 1. In reality, all respondents chose Choice A and therefore, the direction of Set 1 is estimated correctly by SAS. Regarding all choice sets, the goodness-of-fit analysis of the β-values indicates that not all expectations of SAS did correctly representing real scenarios. It is hard to indicate which factors influenced the expectations of SAS most. First, it might be possible that SAS misinterpreted the β-values but that would be very unlikely because the calculations of SAS to estimate the choice probabilities are based on the *assumed* β-values. Therefore, it is more likely that the assumptions of the β-values did not correctly representing real scenarios. A third factor might be that the generalizability of the sample size is not broad enough because the sample size is relatively small and consists mainly of Dutch students. However, it is not yet investigated what happens to the efficiency of the design when the β-values are poorly estimated (Kuhfeld, 2010). If a design appears to have a lower efficiency, the design precision would be less than optimal. Overall, these differences between the expected probabilities and the real percentages proves the exploratory background of this research. Future research could use the actual β-values in order to estimate a better fitting design.

In order to uncover the segments that describe end users of mobile apps and to investigate the characteristics of the segments, a phased approach is used in Latent Gold Choice 5.1. First, the preferences were estimated at individual level (individual part-worth estimation). Second, based on the estimated part-worths, segments were created and third, these segments were profiled by connecting them to the characteristics of the respondents. A latent class analysis integrates these three steps. The latent class analysis assumes that each individual belongs to one segment (Vermunt & Magidson, 2002). Furthermore, all respondents are observed simultaneously so that it is possible to consider all the statistical observations regarding the process of segmenting the respondents which results to more reliable output.

4.3.1 Part-worths estimation

In order to analyze all possible models with the collected data, the software Latent Gold Choice 5.1 allows to choose between inserting only one data file or three separate data files. In this case, it is decided to use the 3-file format because it makes it more easy to define the effects and interactions to be included in a model. A detailed explanation of the analysis in Latent Gold can be found in Appendix VI. The first step of the analysis is by using the aggregate (or 1-class) model because it takes the entire sample into account. Therefore, by using the data of the participants of this survey, the actual utility of each attribute can be identified in the output. In Table 4.2 on the next page, the level of importance of each attribute on an aggregate level can be seen. The z-value provides insight in the significance of each level of influence. If the z-value has a value between -1.65 and 1.65, it indicates that this certain level is *not* significant.

The relative importance of each attribute is obtained as follows:

$$Relative\ Importance = \frac{Max_i - Min_i}{\sum Max_{total} - Min_{total}} * 100$$

Overall, based on the z-values, it can be seen that each attribute is significant. The attribute 'User Rating' is dominating the consumer's choice with a relative importance of 43%, following by the attribute 'Price'. The remaining three attributes have relatively low effects regarding the relative importance.

Table 4.2: Importance Levels of Attributes on an Aggregate Level

Attribute	Influence	z-value	Range	Relative Importance	Rank
Quality Label*				9.23%	4
1) Acknowledged	0.2757	5.6218			
2) Not Acknowledged	0.0880	1.2685	0.6395		
3. No Label	-0.3638	-4.5375			
User Rating*				43.28%	1
1) 5 Stars	1.4278	26.6130			
2) 3 Stars	0.1438	1.9176	2.9994		
3) 1 Star	-1.5716	-18.6486			
Number of Review*				5.14%	5
1) 1 Million	-0.0761	-1.4138			
2) 500.000	0.2162	3.8937	0.2162		
3) 10.000	-0.1401	-2.6646			
App Ranking*				11.30%	3
1) Top 3	0.2662	5.0757			
2) Top 10	0.2508	4.7179	0.7831		
3) Top 50	-0.5169	-8.1162			
Price*				31.05%	2
1) $0.99	0.9664	16.5734			
2) $2.99	0.2187	3.4979	2.1515		
3) $4.99	-1.1851	-15.8768			

1 = best / positive option; 2 = moderate / neutral option; 3 = worst / negative option
Range = highest value of influence – lowest value of influence
*p = < 0.001

4.3.2 Estimation of Number of Segments

After the analysis of the aggregate model, Latent Gold makes it possible to estimate an infinite number of models. Therefore, the best number of segments for this research can be determined. The first output retrieved from Latent Gold is presented in Table 4.3.

The software reports multiple information criteria that weight the fit and the parsimony of a model. The lower the BIC, AIC, AIC3, or CAIC, the better the model. Various researchers (Vermunt & Magidson, 2005; Andrews & Currim, 2003; Dias, 2004), suggest that AIC3 is the best indicator in determining the number of latent classes in choice models. It is important to determine the right number of classes because specifying too few ignores class differences and too many cause the model to be unstable (Vermunt & Magidson, 2005). Table 4.3 shows that a 3- or 4-class model best estimates the provided data. AIC3 (as well as BIC and AIC) provide the lowest values for a 4-class model. Only CAIC has a slightly lower value for a 3-class model. The p-values are excluded as all models have extremely low values and provide no information about which model has the best fit.

When looking back at the aggregate model, it can be seen it has a prediction rate of 74.61%. The prediction table in Appendix VI B shows that this 1-class (aggregate) model correctly predicts 1101 of the 1336 alternative A responses, and 806 of the 953 alternative B responses. Overall, only 1907 of the total 2556 (74.61%) observed choices are predicted correctly. When choosing a 3- or 4-class model, the prediction rate is improved to approximately 84% by accounting for the heterogeneity among the segments. Because the prediction rate is almost equal between these two models, the choice between a 3- or 4-class model is based on the goodness of fit. A 4-class model has a better fit and is therefore chosen over the 3-class model. A 4-class model will provide even further insights into customer behavior by being able to identify four different segments. This is tested by analyzing the 3-class model (Appendix VI E). It is seen that the attribute ‘Number of Reviews’ is not significant in all classes. In addition, solely in class 1, even more values are not significant. Besides, the levels of importance of each attribute showed that the three classes do not differ that much from each other. These classes value mostly the attributes ‘User Rating’ and ‘Price’. This is not the case when choosing a 4-class model, therefore this model could provide other interesting insights.

Table 4.3: Estimation for Goodness of Fit

Model	Class-Choice	LL	BIC(LL)	AIC(LL)	AIC3(LL)	CAIC(LL)	L^2	df	R^2
1	1	-1599,0574	3252,6290	3220,1149	3231,1149	3263,6290	3145,5519	131	0,3916
2	2	-1399,1574	2951,9454	2860,3148	2891,3148	2982,9454	2745,7519	111	0,4855
3	3	-1292,9220	2838,5912	2687,8440	2738,8440	2889,5912	2533,2811	91	0,5634
4	4	-1235,8233	2823,5103	2613,6466	2684,6466	2894,5103	2419,0836	71	0,5865
Refined	4	-1240,8818	2793,9808	2607,7637	2670,7637	2856,9808	2429,2007	79	0.5857

4.3.3 Model Improvements

After selecting the best number of segments, the model should be refined by using restrictions. When analyzing the significance of attributes of the 4-class model, it can be seen that all five attributes are significant (Appendix VI C; p-value < 0.05). This means that all factors significantly influence the decision making of choosing a mobile app. So, none of the attributes have to be excluded by this restriction. Other optimizations through restrictions can be done by investigating the Wald(=) statistic and the z-values. The Wald(=) tests whether regression coefficients are equal between classes. If this statistic shows a significant value, it means that the attribute is segment-specific. The attributes ‘Number of Reviews’ and ‘App Ranking’ have a p-value higher than 0.05. However, when dealing with these attributes by selecting to restrict the effect of these attributes to be class independent, will not lead to great improvements. Besides, some attributes have values between -1.65 and 1.65, which indicates that these estimates are not significant. The class 1, 2, and 3 effects for the attribute ‘Number of Reviews’, as well as class 1 regarding the attribute ‘App Ranking’ are not significant. To deal with these estimates, it is decided to restrict the effect of these classes to zero. The refined model has a better goodness of fit as can be seen in Table 4.3. This model appears to yield the best fit and the analysis was continued with the refined 4-class model.

4.3.4 Attribute Preference per Segment

In the output of Latent Gold 5.1 in Table 4.4, the relative importance of each attribute per segment for the refined 4-class model is provided:

Table 4.4: Importance Levels of Attributes per segment of refined 4-class model

Class 1 (Size: 48.56%) Peer Followers		**Class 2 (Size: 26.92%) Balanced**		**Class 3 (Size: 15.53%) Price-sensitive**		**Class 4 (Size: 9%) Quality Affected**	
Influence	**Relative Importance (Rank)**	**Influence**	**Relative Importance (Rank)**	**Influence**	**Relative Importance (Rank)**	**Influence**	**Relative Importance (Rank)**
Quality Label*	8.17% (3)		12.73% (3)		8.73% (4)		28.02% (1)
1) 0.1710		*0.4998*		*0.4185*		*0.8878*	
2) *-0.3952*		0.1213		0.2823		0.2820	
3) 0.2243		*-0.6211*		*-0.7008*		*-1.1698*	
User Rating*	68.33% (1)		42.54% (1)		24.38% (2)		23.66% (2)
1) *2.6226*		*1.7110*		*0.9745*		*0.6037*	
2) -0.0644		*0.3243*		*1.0758*		*0.5297*	
3) *-2.5581*		*-2.0353*		*-2.0503*		*-1.1334*	
Number of Reviews (p=0.0008)	-		-		-		14.08% (5)
1) 0		0		0		0.2091	
2) 0		0		0		*0.4125*	
3) 0		0		0		*-0.6216*	
App Ranking*	-		12.57% (4)		18.79% (3)		15.30% (4)
1) 0		*0.4201*		*0.9037*		*0.4516*	
2) 0		*0.2669*		*0.6016*		0.2202	
3) 0		*-0.6870*		*-1.5052*		*-0.6718*	
Price*	23.50% (2)		32.16% (2)		48.10% (1)		18.93% (3)
1) *1.0210*		*1.3340*		*2.5673*		*0.5129*	
2) *-0.2604*		0.1640		*1.0315*		0.3642	
3) *-0.7606*		*-1.4979*		*-3.5987*		*-0.8771*	

1 = best / positive option; 2 = moderate / neutral option; 3 = worst / negative option
Values written in *italic* have a significant z-value
*p < 0.001

Table 4.4 clearly shows some differences between each segment. Regarding segment size, the first segment is by far the largest. Furthermore, it can be concluded that the attributes are highly significant. Other differences among the segments can be seen regarding the importance levels. The attribute ‘Number of Reviews’ appeared to be only significant in segment 4, but even in this segment it has a very low importance level. The attribute ‘User rating’ is clearly the most important factor in the first segment, followed by ‘Price’. The second segment, class 2, can be characterized by the fact that ‘User Rating’ as well as ‘Price’ are the most important factors, whereas importance of the attribute ‘Price’ dominates in the third segment. The last segment, class 4, is interesting because the importance levels are less extreme and more equally divided, and the attribute ‘Quality Label’ is most important. In the end, it can be concluded that most customers within the app market focus mainly on the user rating and the price of a certain app. From a managerial perspective, Table 4.4 provide insights on which factors should be targeted by app developers and/or app owners and be able to influence the decision making process of app consumers.

Besides the importance levels of the attributes, Table 4.4 also includes the part-worth utilities of all attribute levels. The value and direction of each attribute level indicates to what degree each level influences the decision making process and which level is most attractive for consumers. Logically, this means that level 1 should have the highest utility and level 3 the lowest or negative utility. To explain this in more detail, take for example segment 1. Level 1 of ‘User Rating’ has the highest and the only positive value. This indicates that the level 1 is the most attractive value of the attribute ‘User Rating’. Level 2 and 3 are lower and negative, indicating it is less attractive than level 1 and negatively influence the decision making process. For segment 1, it is remarkable that these people are slightly more attracted to a scenario where an app has received no quality label at all, instead of a scenario where an app has received a quality label which is acknowledged by the market. For segments 2, 3 and 4 the directions of the influences follow a logical pattern.

Overall, in the literature review it became clear that the quality of apps was often valued from an end user perspective instead of the actual technical quality of the apps. This is supported by the outcome of the Latent Gold analysis in Table 4.4. The decision making process of the majority of the app end users (segment 1 and 2) is mostly influenced by the user rating of apps followed by the price of an app. These segments were not greatly influenced by the attributes ‘Quality Label’ and ‘App Ranking’. Segment 2 is, thus, very similar to segment 1, but distinguishes itself by a higher importance level of the attribute ‘Quality Label’ and it can be seen that the respondents in this segment are more attracted to an app which has a quality label. The end users in segment 3 were most influenced by the price of an app, followed by the user rating. It can be assumed that this segment is likely to use free apps mainly. This is strengthened by the fact that the level 3 influence of price is highly negative. Finally, segment 4 is the only class where the respondents value the attribute ‘Quality Label’ as most important.

4.3.5 Profiling the Segments

With the output of the analysis in Latent Gold 5.1 and the previous steps taken, it is possible to analyze the specific profiles of each segment. The latent class analysis has also taken the covariates into account to link the segments with app end user characteristics. The relative values of each class as well as the significance levels are summarized in Table 4.5.

Table 4.5: Profile of each segment

	Class 1 Peer Followers	Class 2 Balanced	Class 3 Price-sensitive	Class 4 Quality Affected	Characteristics of sample (Appendix V)
Gender (p=0.006)					
Male	82.53%	55.89%	31.23%	52.83%	64.8%
Female	17.47%	44.11%	68.77%	47.17%	35.2%
Age (p=0.54)					
18-24	57.77%	54.91%	50.17%	31.88%	53.5%
25-34	24.34%	35.73%	40.67%	52.07%	34.6%
35-49	9.47%	3.75%	0%	0.09%	5.6%
50+	8.42%	5.61%	9.16%	15.96%	6.3%
Nationality (p=0.15)					
Dutch	80.21%	54.7%	72.94%	67.91%	71.1%
Other European	8.27%	11.3%	13.37%	0.19%	9.2%
North America	2.89%	2.62%	0	7.92%	2.8%
South America	4.31%	13.13%	9.11%	15.82%	8.5%
Asian	4.32%	18.25%	4.57%	8.16%	8.5%
Education (p=0.020)					
High School	34.82%	33.67%	59.25%	0.2%	35.2%
Associate’s Degree	2.89%	2.62%	4.56%	0%	2.8%
Bachelor’s Degree	49.09%	53.42%	32%	68.01%	49.3%
Master’s Degree or Phd	13.20%	10.29%	4.20%	31.79%	12.7%
Profession (p=0.68)					
Part-time	9.82%	8.35%	4.59%	23.83%	9.9%
Full-time	33.04%	20%	31.59%	28.19%	28,9%
Unemployed	2.89%	0%	0%	7.93%	2.1%
Student	52.81%	71.65%	63.83%	40.04%	58.5%
Income (p=0.58)					
Less than €1,500	46.35%	57.03%	45.94%	32.02%	47.9%
€1,500 – €2,999	16.49%	10.94%	18.27%	50.67%	18.3%
€3,000 – €5,999	14.60%	15.50%	17.6%	8.75%	14.8%
€6,000 – €8,999	3.70%	1.13%	0%	0.08%	2.1%
Operating System (p=0.41)					
iOS	34.5%	42.28%	22.9%	39.11%	35.2%
Android	42.94%	54.09%	67.99%	60.88%	51.4%
Other	22.56%	3.63%	9.11%	0.01%	13.4%
Number of Devices (p=0.30)					
1	31.83%	51.06%	58.56%	76.05%	45.1%
2	46.29%	31.39%	27.38%	7.60%	35.9%
3	11.92%	12.22%	14.06%	15.80%	12.7%
More than 3	9.96%	5.33%	0%	0.55%	6.3%

It can be seen that only Gender and Education have p-values smaller than 0.05. This indicates that the difference of these covariates among the four segments are significant. Therefore, it can be concluded that these covariates are the most important characteristics when it comes to describing the profiles of each segment. However, in order to completely profile each segment all covariates will be taken into account.

Starting with segment 1, it can be seen that it mainly consists of Dutch male consumers between the 18 and 34 years old. In addition, most of these consumers have a High School or Bachelor's degree and are currently studying or full-time employed. So, because it consists of many younger people who are still studying, it is not strange that the majority making a relatively lower amount of money. These conclusions are not very surprising when comparing segment 1 with the sample characteristics in Table 4.5, because it is the largest segment and most of the respondents are characterized by these features.

As was concluded in the previous section about the attribute preferences that segment 2 is very similar to segment 1, it can also be said that segment 2 is not very different than segment 1 regarding most of the covariates. A big difference is that segment 2 has a relative equal number of males and females and slightly more foreign respondents. Furthermore, this segment consists mainly of students and make a relatively low amount of money.

Segment 3 consists of predominantly Dutch female consumers between 18 and 34 years old. Again, the majority is still a student, followed by full-time employees. Compared to the previous segments, segment 3 has relatively lower education levels. The relative amount of money made is still very low in this segment.

Just like the differences of the attribute preferences of segment 4 compared to the others, the covariates of this segment are different too. The rate of being a male or female is relatively equal and the majority is Dutch and between the 25-34 years old. Notice that (almost) all respondents in segment 4 are highly educated with a Bachelor's, Master's or PhD degree. 40% of the respondents are student and another 40% are working. The mean income of this segment is also higher than the other three.

Looking at the mobile app related covariates, around 50% of the people in segment 1 have Android and use on average two devices per week. The other half have iOS or various operating systems. Segment 2 is almost equally divided by Android and iOS users, whereby most people use one device per week. Segment 3 and 4, on the other hand, clearly use one Android device in most cases, whereby the last segment includes slightly more people that use iOS as alternative.

As a final remark, it can be seen that some demographics have very low percentages. To test if the segments could be improved by dealing with these factors, it is decided to combine some of the factors which included very low to none respondents. In Appendix VI F the alternative profiles are presented. The changes that have been made are: group of 50+ is combined with the group of 35-49; Associate's Degree is combined with High School; and individuals with an income between €6,000 - €8,999 are combined with the income range of €3,000 - €5,999. However, the outcome is roughly the same as before. Therefore, it is decided to keep the first model.

4.4 Summary

This chapter showed the conjoint analysis and its results with the use of the software Latent Gold 5.1. With these results, a segmentation of the different classes has been provided. First the 1-class (aggregate) model was investigated, where could be seen that all the five attributes were significant. Secondly, through estimation a 4-class model was rendered as it seemed to have the best fit. This 4-class model was further improved by utilizing restrictions. The refined model showed that all five attributes remained significant and indicated which attributes in each segment were most attractive to app consumers. With the importance levels, the segments could be distinguished of each other. Finally, the consumer characteristics were linked to each segment. Gender and Education appeared to be the only significant covariates among the segments.

Chapter V – Conclusion and Discussion

This chapter will pay attention to the conclusions that can be drawn from the literature as well as the data analysis. First the general conclusions will be provided to answer the research questions. After that, the implications, recommendations and the limitations of the research will be presented. Finally, further research will be discussed.

5.1 General Conclusions

The overall goal of this research was to find answers to these two questions: "*How can segments in the direct customer base of Omnext be described in terms of the preferences and characteristics of an independent cloud based quality analysis for mobile apps"* and "*How can segments in the indirect customer base of Omnext be described in terms of the preferences and characteristics of an independent cloud based quality analysis for mobile apps".* In the previous chapter it became clear that the response rate of the direct customer base of Omnext, the developers and owners of apps, was too low and therefore it was not possible to perform a sufficient latent class analysis with the few responses received. That is why conclusions of the first research questions cannot be drawn too. So, only the research questions concerning the indirect customer base of Omnext, the end users, will be discussed.

In order to answer the main problem and its research questions, a latent class analysis with the aid of the software Latent Gold 5.1 has been conducted. The input of the latent class analysis came from 142 respondents which filled out a questionnaire containing eighteen different choice sets. The respondents varied in certain demographics, like gender, age, or education.

1) Preceding the distribution of the questionnaire a literature study has been conducted. Based on the existing literature, it was possible to answer the first two research questions. The first research question stated: *"What are relevant characteristics with which customer segments can be described".*

According to Hamka et al. (2014) the segmentation of the end user segments was typically based on demographic characteristics. In addition, various other studies (Walsh et al., 2010; Plaza et al., 2011; Castells et al., 2004; Head & Ziolkowski, 2012) found that the end users of mobile devices or applications can be segmented based on characteristics as age, nationality, gender, education, and income. Therefore, these characteristics along with other questions about mobile app usage were used for the questionnaire in this research.

2) The second research question, answered in Chapter III, stated: *"What are the most important product attributes for customers when adopting an app".*

Initially, various attributes were derived from literature as well as from interviews with mobile app developers and end users. However, the choice-based conjoint method restricts the number of attributes that can be used. Too many attributes can lead to a situation where respondents will not react in a proper way, but using too few attributes is not recommended because some important characteristics might be excluded (Hair et al., 2010). Eventually, the most important attributes were selected based on (1) previous studies about the different segments existing in the mobile market, (2) the input of two subject matter experts, and (3) a qualitative pre-interview among some end users and mobile app developers: Quality Label, User Rating, Number of Reviews, App Ranking, and Price.

3) Eventually, after the data analysis, the remaining research questions were answered. The third research question stated: *"Do different segments exist in Omnext's indirect customer base on the preferences of an app".*

It was concluded that multiple segments exist within the indirect customer base of Omnext. According to the various information criteria, it appeared that the app market could be segmented into four segments. These four classes have different sizes (48.56%, 26,92%, 15.53%, and 9% respectively) and consists of end user segments with different characteristics.

4) The latent class analysis provided the relative importance of each attribute within each segment. Therefore, it is clearly visible which factors have the most influence on the purchase decision of end users and is it possible to answer the fourth research question, which stated: *"What importance does each of the identified segments attach to the attributes of an app".*

The four segments could be distinguished based on the level of importance of the attributes. The first segment (Peer Followers) valued the 'User Rating' the most with a high importance level of 68.33% followed by the factor 'Price'. In the second segment (Balanced), the 'User Rating' has the highest level of importance again. However, this time with a relative importance of 42.54% it is not extremely high and the factor 'Price' became more important compared to segment 1. The third segment (Price-sensitive) is obviously influenced by the factor 'Price'. The majority of end users in segment 3 valued price the most with a relative importance of 48.10%. As last, the fourth segment (Quality Affected) valued the 'Quality Label' most with a relative importance of 28.02%. This is interesting because in all other three the segments the 'Quality Label' only had a very low influence.

5) Finally, by profiling the segments, Latent Gold 5.1 provided more insights in the segments to answer the fifth research question, which stated: *"With which customer characteristics can each customer segment be described".*

In the refined 4-class model (Appendix VI D) can be seen that only two of the eight customer characteristics are significant. Therefore, gender and the degree of education have an influence on the decision making process of end users. When these two characteristics are linked to the segments, one can conclude that mainly men with a Bachelor's degree are in segment 1, 'Peer followers'. The second segment 'Balanced' consists of approximately the same amount of men and women. The majority of these individuals have a Bachelor's degree (53.42%) or finished High School (33.67%), just like the first segment. The big differences compared with the other segments is that segment 3 'Price-sensitive' consists mainly of women which finished High School. Finally, the fourth segment 'Quality Affected' consists of as many men as women. It is interesting to see that these individuals are all highly educated. 68.01% of the individuals in segment 4 have a Bachelor's degree and 31.79% have a Master's or a PhD degree.

In the end, it can be concluded that the majority in the indirect customer base of Omnext, the end users of apps, are being influenced by the user rating and price of a mobile app when facing a purchase decision. The refined 4-class model showed that a small niche, consisting of highly educated individuals, valued the presence of an acknowledged quality label more than the user rating or the price of an app. Concerning an independent cloud based quality analysis service, the main conclusion is that only a small group of people values high quality of apps and the majority values other peoples' experience based on functionality or are not willing to pay (a certain price) for an app.

5.2 Implications and recommendations

5.2.1 Theoretical implications

The results of this study showed some similarities with previous studies and contributed by investigating the segments within the mobile app market and the perceptions toward the quality of apps in more detail.

1) In the studies of Hamka et al. (2014) and Head & Ziolkowski (2012), end users were segmented based on various characteristics, like age, gender and the usage of mobile apps. Unfortunately, the factors 'age' and 'mobile app usage' were not significant in this study and could not be compared with the previous studies. However, because this research included other factors to segment the customer market, it provided other insights in the mobile app market. Instead of the segmentation based on the amount of usage of mobile apps, this study presented a segmentation based on the purchase or download behavior towards mobile apps. Therefore, developers and owners should be able to adapt their business strategies more effectively, like improving the app based on feedback of the user rating, adapting the pricing strategy, or improving the focus on the technical quality.

2) Furthermore, as was concluded in the study of Kim et al. (2014), this research confirms the fact that user reviews are influential. The user reviews are mainly based on the experience users have when using an app. This research added another visual factor, a quality label, to analyze if it affects the purchase or download behavior of end users. It appeared that for some (highly educated) people the technical quality seemed important and a quality label would make it easily visible if a certain app is 'good'. Kim et al. (2014) also indicated that the perceptions about app price were likely to vary according to individual economic tendencies and the availability of free apps. This study provides further support that economic tendencies likely affect the perception of price. It was seen that lower educated people were more affected by price, whereas highly educated people seemed to choose quality above price. Because many respondents were students, it is likely that lower educated respondents were still studying and did not have a high income. Highly educated people are likely to be older and have a better paying job.

3) In the study of Hildenbrand et al. (2015) it was concluded that the credibility of a quality label is important. This study supports that credibility theory, because in three out of four segments could be seen that the influence of a label that is acknowledged by market is indeed more

positive than a quality label which is not acknowledged by the market. Hildenbrand et al. (2015) also concluded that a quality label is more important when foods are bought. This might be the reason that a quality label seemed not that important regarding mobile apps.

5.2.2 Practical implications

Several managerial implications can be drawn from the conclusions. Some recommendations should be taken into account by app developers and some by software quality analysis service companies, like Omnext.

1) For mobile app developers, the findings imply that they may need to pay close attention to positive user reviews. Kim et al. (2015) already found that the appropriate combination of information and entertainment within apps is likely to be the best fit for end users' needs. However, an app would be perceived as entertaining when it works properly in terms of minimum bugs and crashes. If an app crashes a lot or has other functional errors and people complain about it in their reviews, app developers should be recommended to consider a software quality analysis service because it helps to discover what is wrong in a short time. Of course, app developers might discover defects by themselves, however when an app consists of hundreds or thousands lines of source code, it could be very time consuming. At the time when the app has finally been improved, end users may already found another similar app. A software quality analysis service company should be able to provide example cases to gain developers' trust in the relevant benefits of higher quality software.

2) Though, the original survey meant for app developers and owners could not been analyzed, some personal feedback of developers, the interviews (Appendix I), and an adapted questionnaire which has been distributed to many developers and owners, provided some insights in their interest towards app quality analysis services. It appears that the majority of the interviewees and respondents are willing to use an app quality analysis service if it can be proven that it helps to improve the software maintenance and technical quality. Overall, some developers within relatively larger app development companies were very interested in software quality, whereas some developers within a small app development company did not see some benefit of a software quality analysis. Besides, two interviewees preferred a trial version and one respondent is willing to use an app quality analysis service if he knows how an analysis report looks like and if it is helpful. A trial version could be an effective strategy to present how an

analysis will be performed and to be able to demonstrate how it helps to improve the software quality of an app.

3) In addition, the interviews and the adapted questionnaire provided some basic insights in the importance of various factors. In the questionnaire is asked to rank six factors from 1 (most important) to 6 (least important). The average values are presented in table 5.1:

Table 5.1: Ranking of various factors in an app quality analysis service

Number of projects per year	Number of scans per year	Receiving a quality label	Type of platform	Level of reporting detail	Price of an analysis
3.93	3.64	5.14	4	2.07	2.21

It can be seen that the level of reporting detail is of most importance, followed by the price of an analysis service. Receiving a quality label after a scan, seems not very important. However, this study concluded that a small group of end users values the presence of a quality label relatively high. Besides, some direct feedback of respondents during the distribution of the paper-and-stencil versions, indicated that the function of an app would made them fill out other answers. For simple apps that have many similar alternatives, user rating was most important but when purchasing a crucial app, like a business or banking app, some respondents would value the quality of an app more. Overall, this research can be used to convince the developers and owners of the importance of quality. Not only by being able to present the app quality with a visible label but also by making developers and owners aware that quick improvements and high maintainability of software could have multiple benefits. How to visualize and shaping the quality label is open to discussion.

5.3 Limitations and further research

Research regarding the technical quality of mobile apps is scarce. This study tried to add more insights into the preferences for mobile software quality analysis services regarding app developers and owners as well as app end users. Unfortunately, this study has some limitations that offer opportunities for future research.

First of all, no conclusions could be drawn from the survey meant for app developers and owners because too few respondents filled out the questionnaire. Through feedback of some app developers it became clear that the conjoint analysis structure seemed quite difficult and it took more time to fill out the questionnaire than expected. To include less choice sets would probably not result in higher response rates because many people had opened the survey but only few had finished it and – besides the difficulty of the choice sets – the feedback also made clear that the first pages contained too much text. Further research should re-evaluate the attributes and (1) choose other attributes or (2) use another type of survey or conjoint analysis structure. In this research another type of questionnaire, including questions about the six attributes that were used in the original questionnaire, was already distributed to test if the response rate would increase. Though, the simplicity of the survey was increased, the response rate remained still very low. This could mean that the attributes are vague or that app developers and/or owners are less willing to participate in online surveys. Personal contact might increase the response rate for future research.

Regarding the survey meant for end users, the response rate stayed quite low the first two weeks. Of course, it is normal that online surveys have low response rates but after one week it was noticed that the survey did not work fine on mobile devices. It is assumed that this problem worsened the response rate. Therefore, it is decided to distribute the survey on paper directly in person to possible respondents. This appeared to be successful but resulted in one drawback; a very homogenous sample. The majority of the sample consisted of Dutch students, so the generalizability toward foreign and older end users might be questionable. Eventually, the number of valid responses of 142 is enough to conduct a conjoint analysis, but the sample size is still too small for an actual robust latent class analysis. According to Hair et al. (2010), at least 200 respondents per segment are needed for conducting a thorough segmentation analysis. Therefore, the external validity of this study is limited and the findings need to be viewed with caution. Further research could solve these problems by using a larger sample size and contacting more foreign as well as older end users. This could also result in a higher variety among the segments. Besides, using another survey tool which allows to fill out the survey on mobile devices would also may lead to better responses.

Furthermore, it was seen that only two covariates, gender and education, were significant. The reason for this could be that the sample size mainly consisted of Dutch students. As was seen in the demographic statistics, groups consisting of foreign end users and end users older than 35 were very small and in that case they could not completely represent the real sample. Besides, the factor 'profession' showed that a large part were students, followed by full-time employees. Very few respondents had other profession types. Further research might improve the significance by expending the sample size. This might also provide more insights in the small niche of quality affected people because it is not clear why all respondents in this segment were highly educated. In addition, the attribute ‘number of reviews’ appeared to be significant for only one segment. Excluding this attribute might show other insights in the segments. However, it could also be that respondents interpreted this attribute wrongly, so adapting the definitions might affect the answers but could also affect the response rate. If the text with definitions becomes too long, people would likely drop out the survey too early.

When looking at the complete app market, some more limitations show up. First of all, it may be that not all factors end users would take into account have been included in the study. Besides, the β-values used in SAS to design the questionnaire might have been wrongly predicted. If β-values were wrongly assumed it may have led to a inefficient design and resulted in a distortion of the results. Further research could use the β-values that are based on the findings of this study.

The previous limitations may lead to the discussion about why the user rating is most important and why a quality label has low influences. Most end users could think that apps will not be very harmful in terms of viruses or hackers because 1) they trust the app markets as they say that the apps are analyzed before making them available for public or 2) end users trust the virus scan tools when installed or 3) end users assume that the risk of harmful apps is very low. Possibly, there are more reasons but they cannot be identified from this research. Presenting a visual representation of the quality labels in the questionnaire might have an effect on the choices respondents make. However, how it should be shaped should be analyzed in further research.

Overall, this research is restricted to the data collection and can be improved by gathering more data and by improving the research design. A greater sample will make it possible to conduct a robust latent class analysis.

Bibliography

Atkinson, L. (2013). Smart shoppers? Using QR codes and 'green' smartphone apps to mobilize sustainable consumption in the retail environment. *International Journal of Consumer Studies, 37,* 387-393.

Banker, R.D., Datar, S.M., Kemerer, C.F., & Zweig, D. (1993). Software complexity and maintenance costs. *Communications of the ACM, 36(11),* 81-94.

Baron, J. & Ritov, I. (1994). Reference points and omission bias. *Organizational Behavior & Human Decision Processes, 59,* 475-498

Basili, V.R., Heidrich, J., Lindvall, M., Münch, J., Regardie, M., Rombach, D., Seaman, C., & Trendowicz, A. (2010). Linking software development and business strategy through measurement. *IEEE Computer, 43(4),* 57-65.

Belch, G.E., Belch, M.A., Kerr, G.F., & Powell, I. (2014). *Advertising: an integrated marketing communication perspective*. McGraw-Hill Education.

Best, R.J. (2000). *Market Based Management.* Upper Saddle River N.J.: Prentice Hall.

Bouwman, H., Lopez-Nicolas, C., Molina-Castillo, F. (2008). Explaining mobile commerce services adoption by different type of customers. *Imeti 2008 International MultiConference on Engineering and Technological Innovation, 1*, 358-363.

Burguera, I., Zurutuza, U. & Nadjm-Tehrani, S. Crowdroid: behavior-based malware detection system for Android. *In Proc. of the SPSM*, 2011.

Butler, M. (2011). Android: Changing the Mobile Landscape. *Pervasive Computing, 10(1),* 4-7.

Castells, M., Fernandez-Ardevol, M., Qiu, J., & Sey, A. (2004). *The Mobile Communication Society: A cross-cultural analysis of available evidence on the social uses of wireless communication technology. International Workshop on Wireless Communication Policies and Prospects: A Global Perspective.* University of Southern California, Los Angeles.

Chia, P.H., Yamamoto, Y., & Asokan, N. (2012). Is this App Safe? A Large Scale Study on Application Permissions and Risk Signals. *WWW '12. Proceedings of the 21st international conference on World Wide Web.* 311-320.

Cooper, R.G. (1982). New product success in industrial firms. *Industrial Marketing Management*, *11*(3), 215-223.

Crawford, M. & Di Benedetto, A. (2010). *New Products Management.* McGraw-Hill, 10th edition, international edition, ISBN: 9780073404806

Danneels, E. (1996). Market segmentation: Normative model versus business reality – an exploratory study of apparel retailing in Belgium. *European Journal of Marketing, 30*(6), 36-51.

DeSarbo, W.S., Ramaswamy, V., & Cohen, S.H. (1995). Market segmentation with choice-based conjoint analysis, *Marketing Letters, 6*(2), 37-147.

DeShazo, J., & Fermo, G., (2002). Designing choice sets for stated preference methods: The effects of complexity on choice consistency. *Journal of Environmental Economics and Management. 44,* 123-143.

Dillman, D., Smyth, J. & Christian, L. (2009). *Internet, Mail, and Mixed-Mode Surveys: The Tailored Design Method*, New York: Wiley.

Doyle, P. (2002). *Marketing Management and Strategy.* 3rd ed. Harlow: Prentice Hall.

Engel, J.F., Kollat, D.T. & Blackwell, R.D. (1968). Consumer behavior. Hinsdale, IL: The Dryden Press.

Felt, A.P., Finifter, M., Chin, E., Hanna, S. & Wagner, D. A survey of mobile malware in the wild. *In Proc. of the SPSM*, 2011

Gerlich, R.N., Drumheller, K., & Babb, J. (2015). App Consumption: An Exploratory Analysis of the Uses & Gratifications of Mobile Apps. *Academy of Marketing Studies Journal*, *19*(1), 69-79.

Green, P.E., & Krieger, A. M. (1996). Individualized hybrid models for conjoint analysis. *Management Science*, *42*(6), 850-867.

Green, P.E., & Srinivasan, V. (1978). Conjoint analysis in consumer research: issues and outlook. *Journal of consumer research*, 103-123.

Green, P.E., & Srinivasan, V. (1990). Conjoint analysis in marketing: new developments with implications for research and practice. *The Journal of Marketing*, 3-19.

Haaijer, R., Kamakura, W.A., & Wedel, M. (2001). The 'no-choice' alternative in conjoint choice experiments. *International Journal of Market Research*, *43.*

Hair, J., Black, W., Babin, B., Anderson, R. & Tatham, R. (2010). *Multivariate Data Analysis*. 7th edition, Upper Saddle River, Pearson Prentice Hall.

Hamka, F., Bouwman, H., de Reuver, M., & Kroesen, M. (2014). Mobile customer segmentation based on smartphone measurement. *Telematics and Informatics, 31*(2), 220-227.

Head, M., & Ziolkowski, N. (2012). Understanding student attitudes of mobile phone features: Rethinking adoption through conjoint, cluster and SEM analyses. *Computers and Human Behavior, 18*(6), 2331-2339.

Hildenbrand, A., Kühl, R., & Piper, A. (2015). How Fragile Is the Credibility of a Quality Label? A Qausi-Natural Experiment Using the Example of *Stiftung Warentest. German Association of Agricultural Economists (GEWISOLA), 55th Annual Conference.*

Holzer, A., & Ondrus, J. (2011). Mobile application market: A developer's perspective. *Telematics and Informatics, 28*(1), 22-31.

Huber, J., & Zwerina, K. (1996). The importance of utility balance in efficient choice designs. *Journal of Marketing research*, 307-317.

Inukollu, V.N., Keshamoni, D.D., Kang, T., & Inukolla, M. (2014). Factors influencing quality of mobile apps: role of mobile app development life cycle. *International Journal of Software Engineering & Applications, 5(5),* 15-34.

Jiang, J.J., Klein, G., & Discenza, R. (2002). Perception differences of software success: provider and user views of system metrics. *Journal of Systems and Software, 63(1),* 17-27.

Johnson, R.M., & Orme, B.K. (1996). How many questions should you ask in choice-based conjoint studies. In *ART Forum Proceedings*.

Johnson, G., & Scholes, K. (2002). *Exploring Corporate Strategy: Text and Cases.* Financial Times Prentice Hall.

Jones, C., & Bonsignour, O. (2011). *The economics of software quality*. Addison-Wesley Professional.

Kalafatis, S.P., & Cheston, V. (1997). Normative models and practical applications of segmentation in business markets. *Industrial Marketing Management, 26,* 519-530.

Kim, E., Lin, J.S., & Sung, Y. (2013). To App or Not to App: Engaging Consumers via Branded Mobile Apps. *Journal of Interactive Marketing, 13*(1), 53-65.

Kim, S.C., Yoon, D., & Han, E.K. (2014). Antecedents of mobile app usage among smartphone users. *Journal of Marketing Communications.*

Kolinsky, R. (2014). What is the difference between mobile and classic development? *APRO Software.* Available at: http://mobileapp-development.com/blog/what-is-the-difference-between-mobile-and-classic-development. (Accessed 21 Apr. 2016)

Kotler, P., Keller, K.L., Brady, M., Goodman, M. & Hansen, T. (2012). *Marketing Manage- ment*, Pearson Education Limited, Essex, England, 2nd edition.

Kuhfeld, W.F. (2010). Marketing research methods in SAS. *Experimental Design, Choice, Conjoint, and Graphical Techniques. Cary, NC, SAS-Institute TS-722*.

LaToza, T.D., Towne, W.B., van der Hoek, A, & Herbsleb, J.D. Crowd development. In *Proc. CHASE*, 2013.

Lehman, M.M., Ramil, J.F., Wernick, P.D., Perry, D.E., & Turski, W.M. (1997). Metrics and laws of software evolution – the nineties view. In *Proc. of the Fourth Intl. Software Metrics Symposium (Metrics'97)*, Albuquerque, NM.

Linberg, K. R. (1999). Software developer perceptions about software project failure: a case study. *Journal of Systems and Software*, *49*(2), 177-192.

Louviere, J.J. (1988). Conjoint analysis modelling of stated preferences: a review of theory, methods, recent developments and external validity. *Journal of transport economics and policy*, 93-119.

Maghnati, F., & Ling, K.C. (2013). Exploring the Relationship between Experiential Value and Usage attitude towards Mobile Apps among the Smartphone Users. *International Journal of Business and Management, 8(4).*

Malhotra, N.K. & Peterson, M. (2006). Basic marketing research: a decision-making approach. Prentice Hall.

Manuel, G.C.J., Adrian, S., & Benavent, C. (n.d.) Final report for the course of research methodology.

Maslow, A.H. (1987). *Motivation and Personality*. Pearson Education, Inc. Upper Saddle River, New Jersey, USA, 3rd edition.

Markovich, S., IDC, H., & Moenius, J. (2013). To Hunt or to Scavenge: Competitive Advantage and Competitive Strategy in Platform Industries.

Mens, T, & Tourwé, T. (2004). A Survey of Software Refactoring. *Transactions on software engineering, Vol 30 (2),* 126-139.

Monroe, K.B. (1990). Pricing: Making profitable decisions. McGraw-Hill Pub. Co.

Musson, R., Richards, J., Fisher, D., Bird, C., Bussone, B., & Ganguly, S. Leveraging the crowd: how 48,000 users helped improve lync performance. *IEEE Softw., 30(*4), 2013.

Oestreicher, K. (2009). Segmentation & the Jobs-to-be-done theory: A Conceptual Approach to Explaining Product Failure.

Olson, P. (2014, February 19). Exclusive: The Rags-To-Riches Tale Of How Jan Koum Built WhatsApp Into Facebook's New $19 Billion Baby. *Forbes.* Retrieved from http://www.forbes.com/sites/parmyolson/2014/02/19/exclusive-inside-story-how-jan-koum-built-whatsapp-into-facebooks-new-19-billion-baby/

Orme, B., & King, W.C. (1998). Conducting full-profile conjoint analysis over the Internet. *Quirk's Marketing Research Review*.

Plaza, L., Martin, S., Martin, C., & Medrano, C. (2011). Mobile applications in an aging society: status and trends. *The Journal of Systems and Software, 84*, 1977-1988.

Sadeh, J.L.B.L.N., & Hong, J.I. (2014, July). Modeling users' mobile app privacy preferences: Restoring usability in a sea of permission settings. In *Symposium on Usable Privacy and Security (SOUPS)*.

Sattler, H., & Hensel-börner, S. (2001). A Comparison of Conjoint Measurement with Self-Explicated Approaches. In *Conjoint Measurement* (pp. 121-133). Springer Berlin Heidelberg.

Schach, R. 1999. Software Engineering. Fourth Edition, McGraw-Hill, Boston, MA.

Slaughter, S.A., Harter, D.E., & Krishnan, M.S. (1998) Evaluating the Cost of Software Quality. *Communications of the ACM, 41(8),* 67-73.

Stol, K.J., & Fitzgerald, B. (2014, June). Researching crowdsourcing software development: perspectives and concerns. In *Proceedings of the 1st International Workshop on CrowdSourcing in Software Engineering* (pp. 7-10). ACM.

Thach, E.C., & Olsen, J.E. (2006). Market segment analysis to target young adult wine drinkers. *Agribusiness*, *22*(3), 307.

Urwiler, R., & Frolick, M.N. (2008). The IT value hierarchy: Using Maslow's hierarchy of needs as a metaphor for gauging the maturity level of information technology use within competitive organizations. *Information Systems Management*, *25*(1), 83-88.

Vermeulen, B., Goos, P., & Vandebroek, M. (2008). Models and optimal designs for conjoint choice experiments including a no-choice option. *International Journal of Research in Marketing*, *25*(2), 94-103.

Vermunt, J. K., & Magidson, J. (2002). Latent class cluster analysis. In J. Hagenaars, & A. McCutcheon (Eds.), Applied latent class analysis. (pp. 89-106). Cambridge: Cambridge University Press.

Vermunt, J.K., & Magidson, J. (2005). Latent GOLD 4.0 User's Manual, Statistical Innovations Inc.

Virelli, G. (2001). *Number of levels effect in adaptive and choice-based conjoint analysis* (Doctoral dissertation, Concordia University).

Walsh, S., White, K., & Young, R. (2010). Needing to connect: the impact of self and other on young people's involvement with their mobile phone. *Australian Journal of Psychology, 62*, 194-203.

Wang, H.Y., Liao, C., & Yang, L.H. (2013). What Affects Mobile Application Use? The Roles of Consumption Values. *International Journal of Marketing Studies, 5(2),* 11-22.

Wasserman, A.I. (2010, November). Software engineering issues for mobile application development. In *Proceedings of the FSE/SDP workshop on Future of software engineering research* (pp. 397-400). ACM.

Wind, Y. (1978). Issues and advances in segmentation research. *Journal of Marketing Research, 15*(3), 317-337.

Zikmund, W.G. (1999). *Exploring marketing research*, 8th edition. Chicago: Dryden Press Publishing.

Apple App Review Guidelines. Retrieved from https://developer.apple.com/app-store/review/ on 2015, October 5th.

Meeker, M. (2015, May 26). *Internet Trends 2015: Code Conference.* Retrieved from http:// http://www.kpcb.com/internet-trends.

Pocket Gamer, Application Price Distribution. (2012, September). Retrieved from http://www.pocketgamer.biz/metrics/app-store/app-prices/

Rapid Subversion Adoption Validates Enterprise Readiness and Challenges Traditional Software Configuration Management Leaders. (2007, May 15). Retrieved from http://www.businesswire.com/news/home/20070515005055/en/Rapid-Subversion-Adoption-Validates-Enterprise-Readiness-Challenges.

Statista, Average prices for apps in the Apple App Store as of June 2015 (in U.S. dollars). (2015, June). *Report of Statista about average prices for apps.* Retrieved from http://www.statista.com/statistics/267346/average-apple-app-store-price-app/

VisionMobile, Developer Segmentation. (2014, Q3). *Report of VisionMobile on Developer Segmentation.* Retrieved from http://www.vmob.me/DS3Q14.

Windows Store Policies. Retrieved from https://msdn.microsoft.com/en-us/library/windows/apps/dn764944 on 2015, October 5th.

Appendix

I Interviews: with one manager and two developers within a mobile development company and three app users.

The interviews were held among Dutch individuals and the answers are translated and simplified.

<u>Interviews with the developers and manager</u>

- **On which platforms is your app available?**

M1: iOS, Android, Windows.
D1: Android, iOS, Windows.
D2: iOS, Android, Windows.

- **Do you prefer a 'native', 'hybrid' or 'web based' structure?**

M1: With a simple perspective I prefer hybrid. I choose native when I use a lot of features
D1: I prefer native, but a prototype is more easy with a web-based structure. Native is better in terms of performance and usability. Web-based is slower
D2: I prefer native, but sometimes web-based is more efficient within projects

- **Do you experience a certain behavior regarding the number of users of your mobile app?**

M1: Especially an increase, because mobile devices improve and new platforms arise.
D1: Especially an increase. However, end users have to be convinced to use an app these days. Sometimes a website will still be better than an app.
D2: Especially an increase and this will be for the next 5 years probably. TVs and other devices use apps more often today.

- **Do all your developers work intern or is some development outsourced? Are there advantages in both methods?**

M1: Everyone is working intern, except the developer of the windows edition. Benefit intern: increase knowledge within company. Benefits extern: working more ad hoc in some cases, less pressure on to-do list. However, extern is often more expensive.
D1: Everyone is working intern, except the developer of the windows edition. Planning to employ him intern. Outsourcing is less interesting due to communication issues.
D2: Quite comparable to developer 1.

- **How would you get in contact with an extern developer when you are looking for additional help?**

M1: By using my own network. In some cases, an online market place would be interesting.
D1: First trying by using Twitter in my own network. Secondly, looking at 2nd type connections and thirdly using an online market place.
D2: Checking what other known developers have build. Secondly, using online market places.

- **When testing the apps before making them available for public, are there differences in testing between each platform?**

M1: The differences are visible when testing the users' feedback.
D1: Comparable as M1
D2: It is important that the preference per platform are implemented for each platform. Follow the guidelines of each app store. There are differences between type of testing, such as testing look of feel, usability.

- **What features are tested?**

M1: Mainly usability and slightly the lay-out.
D1: Layout and functionality.
D2: It has to meet certain requirements. Testing usability and quality, but depends on the lines of code or the number of people developing the app.

- **What has the highest costs of developing apps?**

M1: Development! Second, marketing.
D1: Development
D2: Development! Secondly, improvements and design and usability.

- **What results in highest earnings? Customers, advertisements in app, purchase of app, in-app purchases or other?**

M1: Everything when collecting data. Data is earnings. With those data the app can be improved.
D1: In case of games: in-app purchases. Advertisements in app does not lead to high revenues. Besides, a 'pro upgrade' to exclude advertisements or get access to more features might be interesting.
D2: The revenue model depends on the type of app. Main goal is to reach a lot of consumers.

- **Is the quality of your software important when trying to decrease costs? Is quality of interest regarding the end users?**

M1: End users benefit from an app that works as it should be and that has a good performance. 'Good' source code does not have a lot of benefits regarding end users.
D1: Always important when developing an app. It could take more time if quality is low to improve app.
D2: When updating apps, the maintenance is important. When updating an app once per year, it is less important.

- **What is the focus of end users when selecting a mobile app? And on what do you focus when distributing an app? Are there differences?**

M1: Does the app works as it should be. User experience is also important. We slightly focus on source code to detect duplicates.
D1: Usability! Performance. Battery consumption and privacy.
D2: Usability and user interact design are important and not the number of features. The focus of developers is: Trying to distribute the app so it will be used a lot. Sometimes the goal is to increase the number of end users or increase revenues.

- **Will end users be influenced when a quality label is assigned to a 'good' app'? If not, would it then depend on the type of app?**

M1: It indeed depends on the type of app. Large organizations that could experience many risks, like insurance companies, would benefit from a good label. In our case it provides no additional benefits.
D1: It only is interesting when it has a positive brand awareness. Usability and other features have to be good in case the label should be of any interest. Regarding developers, a label might provide some value when selecting external developers.
D2: An independent label could be interesting but it has to be acknowledged by market.

- **How do you prefer to distribute an app? Through an app store, or by using a direct channel?**

M1: App stores
D1: App stores
D2: App stores.

- **Finally, would you experience benefits in using a quality analysis of an independent organization which could help to decrease maintenance costs or decreasing development time?**

M1: When outsourcing some business, a quality analysis would be very interesting.
D1: Yes, but I prefer to start with a trial version to test if such an analysis have indeed benefits. If it proves to be helpful, a contract of the analysis service would be interesting. Unfortunately, presenting the source code to other parties might be have some risks. The more lines of code, the more benefits it has.
D2: Developing a simple app with a small team, would experience not many benefits. Maybe slightly regarding the maintainability of the app. Large companies experience often lower maintainability and therefore a quality analysis might be interesting. Finally, it might improve customer loyalty.

- **Would you prefer to pay per analysis or be subscribed to be able to analyze the app multiple times?**

M1: When developing and updating multiple apps on different platforms, I prefer to be subscribed to the service and be able to use the quality analysis multiple times.
D2: Starting with a free trial to convince potential customers would be best, and afterwards potential customers might be asked to subscribe to the service.

Interviews with three end users

- **What are the top 3 type of apps in your opinion?**

1: Message, photo, and banking
2: Agenda, messages, and email
3: Messages, social media, and email

- **Do you prefer iOS, Android or Windows? Why is that?**

1: Android, no reason.
2: Android because windows it not mature enough and don't like iOS because its Apple
3: iOS, because it works fine and because of brand awareness.

- **Are you willing to pay for an app? If you do, when would you do that? Take for example a banking app or a simple game.**

1: If it has added value to exclude advertisements and to get access to more features. I prefer a trail version in that case first.
2: If it has added value. Mainly same question as person 1.
3: No, I don't use a lot of apps.

- **How would you define the quality of an app? When will an app be 'good' in your opinion?**

1: easy of use and performance. It has to look nice and the battery should last quite long
2: Good performance.
3: Usability.

- **Would quality be a factor to take into account when choosing to download an app?**

1: Quality and privacy is relevant.
2: Yes, for sure. Especially privacy.
3: Yes

- **Would brand awareness be a factor to take into account when choosing to download an app?**

1: Yes
2: Yes
3: Yes

- **Would a quality label, as already can be seen when buying food products, have an influence on your app choice?**

1: No, there are too many quality labels where it is not clear what they present.
2: User reviews are more relevant in my opinion because a quality label will not state if an app is good.
3: No, I look at the user ratings.

- **Would your perception of quality be increased when a quality label is assigned to an app?**

1: It might be interesting.
2: I check mainly the user ratings. Quality label should be acknowledged by market to be interesting.
3: I check almost always the user rating

II Definitions of the selected attributes

<u>Attributes concerning developers and owners</u>

1) **Number of Projects:** The number of projects that could be analyzed each year. A project is defined as one mobile app on one operating system, i.e. Facebook app on Android. The number of projects can be:
a) 1 project
b) 5 projects
c) 10 projects

2) **Number of Scans:** The number of scans that can be performed per project each year. A scan provides insight in the state of the software systems. Multiple scans help to keep track of the improvements of the software systems. The number of scans can be:
a) 4 scans
b) 12 scans
c) More than 12 scans

3) **Level of Reporting Detail:** The level of detail that the final reporting of the analysis will provide. These levels can be:
a) Low: very basic feedback, which only makes clear if the overall maintainability of the software is good, moderate or bad by rating it according a 5-point-scale.
b) Medium: an extension of the basic report. Different factors – like complexity, efficiency, maintainability, et cetera – are rated individually so that it becomes clear which factors are good or which factors need to be improved.
c) High: most detailed report. Includes all ratings on all factors and it presents where the errors are located in the source code, plus suggestions of improvement are made.

4) **Independent Quality Label:** After performing an independent quality analysis on a project, the quality of the mobile application can be indicated by a quality label provided by the company that performs the analysis. The following situations can occur:
a) No Label provided
b) Label provided but **not** acknowledged by market
c) Label provided and acknowledged by market

5) **Type of Platform:** An independent quality analysis of mobile apps could be performed through different independent online platforms. These include:
a) Independent online analysis platform: a cloud service where source code can be analyzed to eventually receive a quality report in return.
b) Version Control System: a platform where source code can be saved, accessed and controlled by authorized persons within a company. During this process an independent quality analysis is maintaining the quality and progression of projects.
c) Online Market place: a communication and business channel (i.e. communicate about and assign mobile app development projects) between (external) developers and/or owners through an online market place where the quality analysis of the source code is provided as an extra service.

6) **Price:** How much you are willing to pay per month in dollars. The price can be:
a) 120 dollars
b) 60 dollars
c) 30 dollars

Attributes concerning end users

1) **User Rating:** The rating of a mobile app according to users in terms of a 5-point-scale, whereby 1 star is 'bad' and 5 stars is 'very good'. The user rating can be:
a) 1 Star
b) 3 Stars
c) 5 Stars

2) **Number of Reviews:** The number of reviews which is displayed in the app stores when selecting a specific app. The number of reviews can be:
a) 10.000 Reviews
b) 500.000 Reviews
c) 1 Million Reviews

3) **App ranking (based on number of downloads):** The app ranking is displayed in the app stores, whereby the most downloaded apps are on top. The app ranking can be:
a) Top 50 (number 11 to 50)
b) Top 10 (number 4 to 10)
c) Top 3 (number 1 to 3)

4) **An Independent Quality Label:** The quality of a mobile app can be indicated by using a quality label. The following situations can occur:
a) No Label provided
b) Label provided and acknowledged by market.
c) Label provided but **not** acknowledged by market.

5) **Price of a mobile application:** The price you are willing to pay for a mobile application. The price can be:
a) 4.99 Dollar
b) 2.99 Dollar
c) 0.99 Dollar

III Development of questionnaire in SAS

Based on the book *Marketing Research Methods in SAS* by Warren F. Kuhfeld (2010), a choice design consists of blocks of several alternatives, and each set of alternatives is called a choice set. These alternatives depend on the levels of the different attributes that are used in a research. The software SAS is very useful to construct various choice sets based on a statistical method. The steps that were taken in SAS are explained in the following:

Development of questionnaire regarding 'Developers' and 'Owners'

Step 1: Determine the number of choice sets

It is first essential to enter in SAS the number of attributes and its respective levels. In the study to analyze the preferences of app 'developers' and 'owners' six attributes (Number of Projects, Number of Scans, Level of Reporting Detail, an Independent Quality Label, Platform, Price of Analysis) are being analyzed. Each of the attributes has three levels. The input

%mktruns (3 3 3 3 3 3);

yields the following output:

```
                              Design Summary
                          Number of
                          Levels       Frequency
                              3             6

              Saturated      = 13
              Full Factorial = 729

              Some Reasonable                      Cannot Be
              Design Sizes       Violations         Divided By
                   18 *              0
                   27 *              0
                   36 *              0
                   15               15                 9
                   21               15                 9
                   24               15                 9
                   30               15                 9
                   33               15                 9
                   13 S             21                 3 9
                   14               21                 3 9

             * - 100% Efficient design can be made with the MktEx macro.
             S - Saturated Design - The smallest design that can be made.
```

The output reveals that a number of 18, 27 or 36 stimuli are good main effects designs and do not lead to any violations. According to Johnson and Orme (1996), there is no degradation of data quality when asking 20 choice tasks. Besides, later tasks are completed much faster by respondents.

So, it can be concluded that 9 or 18 choice sets (27 stimuli can't be divided by two), each consisting of two different stimuli variations, need to be designed for the questionnaire.

Step 2: Creating the Candidate Set

Entering the input

```
%mktex (3 3 3 3 3 3, n=729);
```

yields the following output:

Design Number	D-Efficiency	A-Efficiency	G-Efficiency	Average Prediction Standard Error
1	100.0000	100.0000	100.0000	0.1335

This step is necessary to generate candidate profiles for the survey. It tells SAS that any of the profiles in the full factorial design can be used for any alternative in selecting the choice sets.

Step 3: Defining the Structure of Choice Sets

In this research is decided to give the respondents of the questionnaire choice sets consisting of two different alternatives which they can choose from.

```
%mktlab(data=design, int=f1-f2);
Proc print; run;
```

yields all possible alternatives. In the output of step 1 can be seen that 729 different combinations are possible. In the table on the next page a short representation of the full factorial design including all the combinations are shown where x1, x2, x3, x4, x5, x6 are the attributes. Two columns (f1 and f2) are added containing all 1's that will be used to tell each choice set to have two product alternatives. In this case the 'no preference' option will not be used.

Obs	f1	f2	x1	x2	x3	x4	x5	x6
1	1	1	1	1	1	1	1	1
2	1	1	1	1	1	1	2	2
3	1	1	1	1	1	1	3	3
4	1	1	1	1	1	2	1	1
5	1	1	1	1	1	2	2	2
6	1	1	1	1	1	2	3	3
7	1	1	1	1	1	3	1	1
8	1	1	1	1	1	3	2	2
9	1	1	1	1	1	3	3	3
10	1	1	1	1	2	1	1	1
11	1	1	1	1	2	1	2	2

. . .

719	1	1	3	3	2	3	2	2
720	1	1	3	3	2	3	3	3
721	1	1	3	3	3	1	1	1
722	1	1	3	3	3	1	2	2
723	1	1	3	3	3	1	3	3
724	1	1	3	3	3	2	1	1
725	1	1	3	3	3	2	2	2
726	1	1	3	3	3	2	3	3
727	1	1	3	3	3	3	1	1
728	1	1	3	3	3	3	2	2
729	1	1	3	3	3	3	3	3

Step 4: Creating the Choice Design

In the last step, all previously steps will be combined in order to create the desired number of choice sets. As is discussed in Chapter III, four requirements characterize efficient choice designs. By using the %MktEx macro, it is possible to generate an efficient design. Additionally, the assumed β-values derived in Chapter III, are included in the input. SAS treats the last attribute level as a benchmark level, which means the value 0 is automatically assigned to this level and it does not need to be entered. Therefore, the following input is used:

```
%mktex(3 3 3 3 3 3, n=729);
%mktlab(data=design, int=f1-f2)
%choiceff(data=final, model=class(x1-x6), nsets=18, maxiter=20, seed=17, flags=f1-f2, beta=-1 1 -1 1
-1 1 -0.66 0.66 -0.33 0.33 -1 1);
proc print; by set; id set; run;
```

Table A1: Statistical values of Design Questionnaire 'Developers and Owners'

Attribute	n	Variable	Label	Variance	Beta	DF	Standard Error
Number of Projects	1	x10	5				
	2	x11	1	1.01904	-1	1	1.00948
	3	x12	10	1.09212	1	1	1.04505
Number of Scans	4	x20	12				
	5	x21	4	1.08065	-1	1	1.03954
	6	x22	12+	1.07116	1	1	1.03497
Level of Reporting Detail	7	x30	Medium				
	8	x31	Low	1.10515	-1	1	1.05126
	9	x32	High	1.07068	1	1	1.03474
Independent Quality Label	10	x40	Label unknown				
	11	x41	No Label	0.77332	-0.66	1	0.87939
	12	x42	Label known	0.75642	0.66	1	0.86972
Platform	13	x50	Extended				
	14	x51	Basic	0.58727	-0.33	1	0.76634
	15	x52	Complete	0.53917	0.33	1	0.74328
Price	16	x60	60				
	17	x61	120	1.14474	-1	1	1.06993
	18	x62	30	1.05874	1	1	1.02895

The output (Table A1) is slightly adapted to make it more readable. It indicates that the different sets are optimally designed based on multiple combinations SAS has calculated. For an optimal survey the variance terms will be mostly equal and the degrees of freedom (DF) should be 1 for each. According to this output it is concluded that 18 choice sets are needed. The same statistical analysis was made to test if it was possible to make an optimal design with 9 choice sets, but the variances were very different and some variables had a DF of 0.

The second part of the output shows an optimal choice design consisting of 18 choice sets.

Set	Design	Efficiency	Index	Prob	n	f1	f2	x1	x2	x3	x4	x5	x6
1	13	1.90217	621	0.50250	433	1	1	3	2	2	3	3	3
	13	1.90217	311	0.49750	434	1	1	2	1	3	2	2	2
Set	**Design**	**Efficiency**	**Index**	**Prob**	**n**	**f1**	**f2**	**x1**	**x2**	**x3**	**x4**	**x5**	**x6**
2	13	1.90217	559	0.26894	435	1	1	3	1	3	3	1	1
	13	1.90217	167	0.73106	436	1	1	1	3	1	1	2	2
Set	**Design**	**Efficiency**	**Index**	**Prob**	**n**	**f1**	**f2**	**x1**	**x2**	**x3**	**x4**	**x5**	**x6**
3	13	1.90217	403	0.66150	437	1	1	2	2	3	3	3	1
	13	1.90217	201	0.33850	438	1	1	1	3	2	2	1	3
Set	**Design**	**Efficiency**	**Index**	**Prob**	**n**	**f1**	**f2**	**x1**	**x2**	**x3**	**x4**	**x5**	**x6**
4	13	1.90217	483	0.41824	439	1	1	2	3	3	3	2	3
	13	1.90217	584	0.58176	440	1	1	3	2	1	2	3	2

Set	Design	Efficiency	Index	Prob	n	f1	f2	x1	x2	x3	x4	x5	x6
5	13	1.90217	661	0.26503	441	1	1	3	3	1	2	2	1
	13	1.90217	111	0.73497	442	1	1	1	2	2	1	1	3
Set	**Design**	**Efficiency**	**Index**	**Prob**	**n**	**f1**	**f2**	**x1**	**x2**	**x3**	**x4**	**x5**	**x6**
6	13	1.90217	40	0.33626	443	1	1	1	1	2	2	2	1
	13	1.90217	668	0.66374	444	1	1	3	3	1	3	1	2
Set	**Design**	**Efficiency**	**Index**	**Prob**	**n**	**f1**	**f2**	**x1**	**x2**	**x3**	**x4**	**x5**	**x6**
7	13	1.90217	109	0.42068	445	1	1	1	2	2	1	1	1
	13	1.90217	564	0.57932	446	1	1	3	1	3	3	2	3
Set	**Design**	**Efficiency**	**Index**	**Prob**	**n**	**f1**	**f2**	**x1**	**x2**	**x3**	**x4**	**x5**	**x6**
8	13	1.90217	683	0.50250	447	1	1	3	3	2	1	3	2
	13	1.90217	336	0.49750	448	1	1	2	2	1	2	1	3
Set	**Design**	**Efficiency**	**Index**	**Prob**	**n**	**f1**	**f2**	**x1**	**x2**	**x3**	**x4**	**x5**	**x6**
9	13	1.90217	187	0.26698	449	1	1	1	3	1	3	3	1
	13	1.90217	543	0.73302	450	1	1	3	1	3	1	1	3
Set	**Design**	**Efficiency**	**Index**	**Prob**	**n**	**f1**	**f2**	**x1**	**x2**	**x3**	**x4**	**x5**	**x6**
10	13	1.90217	261	0.49750	451	1	1	2	1	1	2	3	3
	13	1.90217	625	0.50250	452	1	1	3	2	3	1	2	1
Set	**Design**	**Efficiency**	**Index**	**Prob**	**n**	**f1**	**f2**	**x1**	**x2**	**x3**	**x4**	**x5**	**x6**
11	13	1.90217	537	0.41824	453	1	1	3	1	2	3	2	3
	13	1.90217	475	0.58176	454	1	1	2	3	3	2	3	1
Set	**Design**	**Efficiency**	**Index**	**Prob**	**n**	**f1**	**f2**	**x1**	**x2**	**x3**	**x4**	**x5**	**x6**
12	13	1.90217	580	0.72909	455	1	1	3	2	1	2	2	1
	13	1.90217	80	0.27091	456	1	1	1	1	3	3	3	2
Set	**Design**	**Efficiency**	**Index**	**Prob**	**n**	**f1**	**f2**	**x1**	**x2**	**x3**	**x4**	**x5**	**x6**
13	13	1.90217	235	0.58176	457	1	1	1	3	3	3	1	1
	13	1.90217	495	0.41824	458	1	1	3	1	1	1	3	3
Set	**Design**	**Efficiency**	**Index**	**Prob**	**n**	**f1**	**f2**	**x1**	**x2**	**x3**	**x4**	**x5**	**x6**
14	13	1.90217	454	0.66150	459	1	1	2	3	2	3	2	1
	13	1.90217	153	0.33850	460	1	1	1	2	3	2	3	3
Set	**Design**	**Efficiency**	**Index**	**Prob**	**n**	**f1**	**f2**	**x1**	**x2**	**x3**	**x4**	**x5**	**x6**
15	13	1.90217	234	0.65701	461	1	1	1	3	3	2	3	3
	13	1.90217	245	0.34299	462	1	1	2	1	1	1	1	2
Set	**Design**	**Efficiency**	**Index**	**Prob**	**n**	**f1**	**f2**	**x1**	**x2**	**x3**	**x4**	**x5**	**x6**
16	13	1.90217	705	0.27091	463	1	1	3	3	3	1	1	3
	13	1.90217	295	0.72909	464	1	1	2	1	2	3	3	1
Set	**Design**	**Efficiency**	**Index**	**Prob**	**n**	**f1**	**f2**	**x1**	**x2**	**x3**	**x4**	**x5**	**x6**
17	13	1.90217	330	0.34074	465	1	1	2	2	1	1	2	3
	13	1.90217	524	0.65926	466	1	1	3	1	2	2	1	2
Set	**Design**	**Efficiency**	**Index**	**Prob**	**n**	**f1**	**f2**	**x1**	**x2**	**x3**	**x4**	**x5**	**x6**
18	13	1.90217	104	0.49750	467	1	1	1	2	1	3	2	2
	13	1.90217	439	0.50250	468	1	1	2	3	2	1	3	1

Development of questionnaire regarding 'End users'

Step 1: Determine the number of choice sets

Again, it is essential to enter in SAS the number of attributes and its respective levels. In the study to analyze the preferences of app 'end users' five attributes (Performance Relevance, Privacy Relevance, Security Relevance, an Independent Quality Label, Price of an app) are being analyzed. Each of the attributes has three levels. The input

%mktruns (3 3 3 3 3);

yields the following output:

```
                         Design Summary
                      Number of
                       Levels       Frequency
                          3             5

            Saturated      = 11
            Full Factorial = 243

            Some Reasonable                     Cannot Be
            Design Sizes       Violations        Divided By
                 18 *              0
                 27 *              0
                 36 *              0
                 12               10               9
                 15               10               9
                 21               10               9
                 24               10               9
                 30               10               9
                 33               20               9
                 11 S             15               3  9

           * - 100% Efficient design can be made with the MktEx macro.
           S - Saturated Design - The smallest design that can be made.
               Note that the saturated design is not one of the
               recommended designs for this problem.  It is shown
               to provide some context for the recommended sizes.
```

The output reveals that a number of 18, 27 or 36 stimuli are good main effects designs and do not lead to any violations. Like the previous case, it can be concluded that 9 or 18 choice sets, each consisting of two different stimuli variations, need to be designed for the questionnaire.

Step 2: Creating the Candidate set

Entering the input

%mktex (3 3 3 3 3, n=243);

shows the following output:

Design Number	D-Efficiency	A-Efficiency	G-Efficiency	Average Prediction Standard Error
1	100.0000	100.0000	100.0000	0.2128

Step 3: Defining the Structure of Choice Sets

In this research is decided to give the respondents of the questionnaire choice sets consisting of two different alternatives which they can choose from.

%mktlab(data=design, int=f1-f2);
Proc print; run;

yields all possible alternatives. In the output of step 1 can be seen that 243 different combinations are possible. In the following table a short representation of the full factorial design including all the combinations are shown where x1, x2, x3, x4, x5 are the attributes. In this case the 'no preference' option will be used, but will be added manually later.

Obs	f1	f2	x1	x2	x3	x4	x5
1	1	1	1	1	1	1	1
2	1	1	1	1	1	1	2
3	1	1	1	1	1	1	3
4	1	1	1	1	1	2	1
5	1	1	1	1	1	2	2
6	1	1	1	1	1	2	3
7	1	1	1	1	1	3	1
8	1	1	1	1	1	3	2
9	1	1	1	1	1	3	3
10	1	1	1	1	2	1	1
11	1	1	1	1	2	1	2

. . .

233	1	1	3	3	2	3	2
234	1	1	3	3	2	3	3
235	1	1	3	3	3	1	1
236	1	1	3	3	3	1	2
237	1	1	3	3	3	1	3
238	1	1	3	3	3	2	1
239	1	1	3	3	3	2	2
240	1	1	3	3	3	2	3
241	1	1	3	3	3	3	1
242	1	1	3	3	3	3	2
243	1	1	3	3	3	3	3

Step 4: Creating the Choice Design

Just like the research regarding 'Developers and Owners', the %MktEx macro is used again to generate an efficient design. Additionally, the assumed β-values derived in Chapter 3, are included in the input. SAS treats the last attribute level as a benchmark level, which means the value 0 is automatically assigned to this level and it does not need to be entered. Therefore, the following input is used:

```
%mktex(3 3 3 3 3, n=243);
%mktlab(data=design, int=f1-f2)
%choiceff(data=final, model=class(x1-x5), nsets=18, maxiter=20, seed=17, flags=f1-f2, beta=-1 1 -
0.33 0.33 -0.33 0.33 -0.66 0.66 -0.66 0.66);
proc print; by set; id set; run;
```

Table A2: Statistical value of questionnaire design 'End users'

Attribute	n	Variable	Label	Variance	Beta	DF	Standard Error
User Rating	1	x10	3				
	2	x11	1	1.01469	-1	1	1.00732
	3	x12	5	0.92909	1	1	0.96289
Number of Reviews	4	x20	500.000				
	5	x21	10.000	0.58694	-0.33	1	0.76612
	6	x22	1 million	0.56099	0.33	1	0.74899
App ranking based on #downloads	7	x30	Top 10				
	8	x31	Top 50	0.54787	-0.33	1	0.74018
	9	x32	Top 3	0.56988	0.33	1	0.75490
Independent Quality Label	10	x40	Label unknown				
	11	x41	No Label	0.78021	-0.66	1	0.88330
	12	x42	Label known	0.76011	0.66	1	0.87184
Price	13	x50	Medium				
	14	x51	Free	0.70603	-0.66	1	0.84026
	15	x52	High	0.78554	0.66	1	0.88631

The output (Table A2) is slightly adapted to make it more readable. It indicates that the different sets are optimally designed based on multiple combinations SAS has calculated. In the case of the questionnaire intended for 'End users', 18 choice sets are needed too. The same statistical analysis was made to test if it was possible to make an optimal design with 9 choice sets, but the variances were very different and some variables had a DF of 0.

Below, the questionnaire intended for 'End users' is presented.

Set	Design	Efficiency	Index	Prob	n	f1	f2	x1	x2	x3	x4	x5
1	1	1.98376	216	0.72711	1	1	1	3	2	3	3	3
	1	1.98376	136	0.27289	2	1	1	2	3	1	1	1
Set	**Design**	**Efficiency**	**Index**	**Prob**	**n**	**f1**	**f2**	**x1**	**x2**	**x3**	**x4**	**x5**
2	1	1.98376	83	0.42068	3	1	1	2	1	1	1	2
	1	1.98376	240	0.57932	4	1	1	3	3	3	2	3
Set	**Design**	**Efficiency**	**Index**	**Prob**	**n**	**f1**	**f2**	**x1**	**x2**	**x3**	**x4**	**x5**
3	1	1.98376	47	0.49750	5	1	1	1	2	3	1	2
	1	1.98376	178	0.50250	6	1	1	3	1	2	3	1
Set	**Design**	**Efficiency**	**Index**	**Prob**	**n**	**f1**	**f2**	**x1**	**x2**	**x3**	**x4**	**x5**
4	1	1.98376	93	0.58662	7	1	1	2	1	2	1	3
	1	1.98376	59	0.41338	8	1	1	1	3	1	2	2
Set	**Design**	**Efficiency**	**Index**	**Prob**	**n**	**f1**	**f2**	**x1**	**x2**	**x3**	**x4**	**x5**
5	1	1.98376	185	0.49750	9	1	1	3	1	3	2	2
	1	1.98376	124	0.50250	10	1	1	2	2	2	3	1
Set	**Design**	**Efficiency**	**Index**	**Prob**	**n**	**f1**	**f2**	**x1**	**x2**	**x3**	**x4**	**x5**
6	1	1.98376	150	0.73106	11	1	1	2	3	2	2	3
	1	1.98376	215	0.26894	12	1	1	3	2	3	3	2
Set	**Design**	**Efficiency**	**Index**	**Prob**	**n**	**f1**	**f2**	**x1**	**x2**	**x3**	**x4**	**x5**
7	1	1.98376	219	0.34299	13	1	1	3	3	1	1	3
	1	1.98376	17	0.65701	14	1	1	1	1	2	3	2
Set	**Design**	**Efficiency**	**Index**	**Prob**	**n**	**f1**	**f2**	**x1**	**x2**	**x3**	**x4**	**x5**
8	1	1.98376	227	0.41581	15	1	1	3	3	2	1	2
	1	1.98376	103	0.58419	16	1	1	2	1	3	2	1
Set	**Design**	**Efficiency**	**Index**	**Prob**	**n**	**f1**	**f2**	**x1**	**x2**	**x3**	**x4**	**x5**
9	1	1.98376	49	0.49750	17	1	1	1	2	3	2	1
	1	1.98376	171	0.50250	18	1	1	3	1	1	3	3
Set	**Design**	**Efficiency**	**Index**	**Prob**	**n**	**f1**	**f2**	**x1**	**x2**	**x3**	**x4**	**x5**
10	1	1.98376	181	0.42068	19	1	1	3	1	3	1	1
	1	1.98376	63	0.57932	20	1	1	1	3	1	3	3
Set	**Design**	**Efficiency**	**Index**	**Prob**	**n**	**f1**	**f2**	**x1**	**x2**	**x3**	**x4**	**x5**
11	1	1.98376	201	0.41581	21	1	1	3	2	2	1	3
	1	1.98376	160	0.58419	22	1	1	2	3	3	3	1
Set	**Design**	**Efficiency**	**Index**	**Prob**	**n**	**f1**	**f2**	**x1**	**x2**	**x3**	**x4**	**x5**
12	1	1.98376	33	0.57932	23	1	1	1	2	1	2	3
	1	1.98376	241	0.42068	24	1	1	3	3	3	3	1

Set	Design	Efficiency	Index	Prob	n	f1	f2	x1	x2	x3	x4	x5
13	1	1.98376	193	0.41581	25	1	1	3	2	1	2	1
	1	1.98376	156	0.58419	26	1	1	2	3	3	1	3
Set	**Design**	**Efficiency**	**Index**	**Prob**	**n**	**f1**	**f2**	**x1**	**x2**	**x3**	**x4**	**x5**
14	1	1.98376	135	0.42068	27	1	1	2	2	3	3	3
	1	1.98376	230	0.57932	28	1	1	3	3	2	2	2
Set	**Design**	**Efficiency**	**Index**	**Prob**	**n**	**f1**	**f2**	**x1**	**x2**	**x3**	**x4**	**x5**
15	1	1.98376	83	0.79249	29	1	1	2	1	1	1	2
	1	1.98376	43	0.20751	30	1	1	1	2	2	3	1
Set	**Design**	**Efficiency**	**Index**	**Prob**	**n**	**f1**	**f2**	**x1**	**x2**	**x3**	**x4**	**x5**
16	1	1.98376	17	0.41581	31	1	1	1	1	2	3	2
	1	1.98376	193	0.58419	32	1	1	3	2	1	2	1
Set	**Design**	**Efficiency**	**Index**	**Prob**	**n**	**f1**	**f2**	**x1**	**x2**	**x3**	**x4**	**x5**
17	1	1.98376	118	0.79249	33	1	1	2	2	2	1	1
	1	1.98376	6	0.20751	34	1	1	1	1	1	2	3
Set	**Design**	**Efficiency**	**Index**	**Prob**	**n**	**f1**	**f2**	**x1**	**x2**	**x3**	**x4**	**x5**
18	1	1.98376	197	0.73106	35	1	1	3	2	1	3	2
	1	1.98376	15	0.26894	36	1	1	1	1	2	2	3

IV Final Version of Survey

Final version intended for 'Developers' and 'Owners'

Thank you for helping me with my Master's thesis at University of Technology Eindhoven by taking part in this survey. Your response is appreciated as it can not only assist me to graduate but also helps to create a software quality service that might increase the overall quality of mobile apps. I am currently studying MSc Innovation Management, and the topic of this thesis focuses on the behavior of mobile app developers and owners (i.e. mobile app companies or people who do not develop apps themselves but have them built by others) with respect to an independent mobile quality analysis. All collected data will be treated anonymously and will be deleted after finalizing the project.

The survey should only take 5 – 10 minutes to complete. During the main part of this survey, you will see a few sets of different independent mobile quality analysis scenarios. I would like to ask you to read all provided information carefully and only think briefly about your responses.

If you have any questions, please feel free to contact me.

With kind regards,

Mike Loeffen
m.t.c.loeffen@student.tue.nl

*For participating in this survey I will donate €1 to a charitable organization of your choice.

NEXT PAGE

On the following pages you will see 18 scenarios with various independent quality analysis services for mobile apps. Each time you are asked to choose between two scenarios which you prefer most. All the packages consist of the same six attributes and each of these attributes has three different levels.

An independent quality analysis is defined as an automated analysis which identifies errors, risks and other issues regarding the maintainability and technical quality of your mobile application based on the ISO-25010 standard.

In the following an explanation of the attributes with the associated levels is provided:

1) **Number of Projects:** The number of projects that could be analyzed each year. A project is defined as one mobile app on one operating system, i.e. Facebook app on Android. The number of projects can be:
a) 1 project
b) 5 projects
c) 10 projects

2) **Number of Scans:** The number of scans that can be performed per project each year. A scan provides insight in the state of the software systems. Multiple scans help to keep track of the improvements of the software systems. The number of scans can be:
a) 4 scans
b) 12 scans
c) More than 12 scans

3) **Level of Reporting Detail:** The level of detail that the final reporting of the analysis will provide. These levels can be:
a) Low: very basic feedback, which only makes clear if the overall maintainability of the software is good, moderate or bad by rating it according a 5-point-scale.
b) Medium: an extension of the basic report. Different factors – like complexity, efficiency, maintainability, et cetera – are rated individually so that it becomes clear which factors are good or which factors need to be improved.
c) High: most detailed report. Includes all ratings on all factors and it presents where the errors are located in the source code, plus suggestions of improvement are made.

4) **Independent Quality Label:** After performing an independent quality analysis on a project, the quality of the mobile application can be indicated by a quality label provided by the company that performs the analysis. The following situations can occur:
a) No Label provided
b) Label provided but **not** acknowledged by market
c) Label provided and acknowledged by market

5) **Type of Platform:** An independent quality analysis of mobile apps could be performed through different independent online platforms. These include:
a) Independent online analysis platform: a cloud service where source code can be analyzed to eventually receive a quality report in return.
b) Version Control System: a platform where source code can be saved, accessed and controlled by authorized persons within a company. During this process an independent quality analysis is maintaining the quality and progression of projects.
c) Online Market place: a communication and business channel (i.e. communicate about and assign mobile app development projects) between (external) developers and/or owners through an online market place where the quality analysis of the source code is provided as an extra service.

6) **Price:** How much you are willing to pay per month in dollars. The price can be:
a) 120 dollars
b) 60 dollars
c) 30 dollars

NEXT PAGE

Set 1: Which independent quality analysis of mobile apps do you prefer most?

<table>
<tr><th></th><th>Option A</th><th>Option B</th></tr>
<tr><td>Number of Projects (/year)</td><td>10 Projects</td><td>5 Projects</td></tr>
<tr><td>Number of Scans (/year)</td><td>12 Scans</td><td>4 Scans</td></tr>
<tr><td>Level of Reporting Detail</td><td>Medium

</td><td>High

</td></tr>
<tr><td>Independent Quality Label</td><td>Label provided and acknowledged by market</td><td>Label provided but <u>not</u> acknowledged by market</td></tr>
<tr><td>Type of Platform</td><td>Online Market Place</td><td>Version Control System</td></tr>
<tr><td>Price (/month)</td><td>30 dollars</td><td>60 dollars</td></tr>
</table>

**Only the first question is presented in the appendix to give an idea how the choice sets looked like. The other choice sets were very similar to the example above.*

NEXT PAGE

What is your gender?

- Male
- Female

What is your age?

What is your nationality?

What is the highest level of education you have completed?

- No degree
- High School
- Associate degree
- Bachelor's degree
- Master's degree
- Doctor's degree (PhD) or postgraduate degree
- Other

How are you related to the mobile app market?

- I am building apps myself (app developer)
- I am not building apps myself but have them built by someone else (owner of a mobile app)
- Both

What is your goal of the mobile app development?

- Growing your Business
- Seeking revenues
- Just a hobby
- Combination of above options

Finally, would you eventually be willing to use an independent quality analysis for mobile apps, if it helps you to improve the software maintenance and technical quality and to reduce costs and risks?

- Yes
- No
- Depends: ___

If you are interested to be informed about future developments regarding the concept of an online independent quality analysis, then please enter your email address in the box below:

NEXT PAGE

Thank you for your participation.

If you know other mobile app developers and/or owners who would like to help me by taking this survey, then please **<click here>.**

If you have additional questions about this survey, please email m.t.c.loeffen@student.tue.nl

Final version intended for 'End users'

Thank you for helping me with my Master's thesis at Eindhoven University by taking part in this survey. Your response is appreciated as it can not only assist me to graduate but also helps to create a software quality service that might increase the overall quality of mobile apps. I am currently studying MSc Innovation Management, and the topic of this thesis focuses on the behavior of mobile app users with respect to the quality of mobile apps. All collected data will be treated anonymously and will be deleted after finalizing the project.

The survey should only take 5 – 10 minutes to complete. During the main part of this survey, you will see a few sets of different mobile app (review) characteristics. I would like to ask you to read all provided information carefully and only think briefly about your responses.

If you have any questions, please feel free to contact me.
Please click 'Next' to start the survey.

With kind regards,

Mike Loeffen
m.t.c.loeffen@student.tue.nl

*For participating in this survey I will donate €1 to a charitable organization of your choice.

NEXT PAGE

On the following 18 pages you will see various characteristics that are often considered before deciding to download a mobile app. Each time you are asked to choose between two scenarios which you prefer most. All the scenarios consist of the same five attributes and each of these attributes has three different levels.

In the following an explanation of the attributes with the associated levels is provided. Assume that the **function (i.e. financial, business, gaming, et cetera) of the app is equal** in each set:

1) **An Independent Quality Label:** A quality label assigned to a mobile app could provide an indication about the quality of an app. The following situations can occur:
a) No Label provided
b) Label provided and acknowledged by market.
c) Label provided but **not** acknowledged by market.

2) **Number of Reviews:** The mobile apps in the app stores are often organized based on various factors, whereby the most downloaded are on top of the page. The number of reviews can be:
a) 10.000 Reviews
b) 500.000 Reviews
c) 1 Million Reviews

3) **App ranking (based on number of downloads):** The app ranking is displayed in the app stores, whereby the most downloaded apps are on top. The apps can be assigned to the following levels:
a) Top 50 (*app is positioned between number 11 to 50*)
b) Top 10 (*app is positioned number 4 to 10*)
c) Top 3 (*app is positioned number 1 to 3*)

4) **User Rating:** The rating of a mobile app as can be seen in the app stores, is presented by a 5-point-scale; 1 star refers to a 'bad' app and 5 stars refers to to an 'excellent' app. The user rating can be:
a) 1 Star
b) 3 Stars
c) 5 Stars

5) **Price of a mobile application:** The price you are willing to pay for a mobile application. The price can be:
a) 4.99 Dollar
b) 2.99 Dollar
c) 0.99 Dollar

NEXT PAGE

Set 1: Which mobile app alternative do you prefer?

	Option A	**Option B**
Independent Quality Label	Label provided and acknowledged by market	No label provided
User Rating	5 Stars	3 Stars
Number of Reviews	500.000 Reviews	1 million Reviews
App Ranking (based on number of downloads)	Top 3	Top 50
Price of an app	$0.99	$4.99
	No Preference	

**Only the first question is presented in the appendix to give an idea how the choice sets looked like. The other choice sets were very similar to the example above.*

NEXT PAGE

What is your gender?

- Male
- Female

What is your age?

What is your Nationality?

Which mobile Operating System are you using?

- Apple iOS
- Google Android
- Microsoft Windows
- Other

How many different mobile devices are you using weekly?

- 1
- 2
- 3
- More than 3

What is the highest level of education you have completed?

- No degree
- High School
- Associate degree
- Bachelor's degree
- Master's degree
- Doctor's degree (PhD) or postgraduate degree
- Other

What is your current job situation?

- Part-time employee
- Full-time employee
- Unemployed
- Student
- I prefer not to answer

What is your aggregated monthly household income after taxes?

- Less than 1.500€
- 1.500€ - 2,999€
- 3,000€ - 5.999€
- 6,000€ - 8,999€
- 9,000€ or more
- I prefer not to answer

If you are interested in the results of the study, then please enter your email address in the box below:

NEXT PAGE

Thank you for your participation.

If you know other mobile app developers and/or owners who would like to help me by taking this survey, then please **<click here>.**

If you have additional questions, pleas email to m.t.c.loeffen@student.tue.nl

V Sample Descriptive Statistics

Age		
	Mean	27.2
	Standard deviation	9.9
Gender		
	Male	64.8%
	Female	35.2%
Nationality		
	Dutch	71.1%
	Other European	9.2%
	North-American	2.8%
	South-American	8.5%
	Asian	8.5%
Education		
	High School	35.2%
	Associate's Degree	2.8%
	Bachelor's Degree	49.3%
	Master's Degree	8.5%
	Doctor's Degree	2.8%
	Other	1.4%
Current profession		
	Part-time Employee	9.9%
	Full-time Employee	28,9%
	Unemployed	2.1%
	Student	58.5%
	No answer	0.7%
Income		
	<€1500	47.9%
	€1500-€2999	18.3%
	€3000-€5999	14.8%
	€6000-€8999	2.1%
	No answer	16.9%
Operating System		
	iOS	35.2%
	Android	51.4%
	Windows	4.2%
	iOS and Android	7%
	iOS and Windows	1.4%
	Android and Windows	0.7%
Number of Devices		
	1	45.1%
	2	35.9%
	3	12.7%
	More than 3	6.3%

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Zscore: Nationality	152	-,55572	2,46102	,0000000	1,00000000
Zscore: Age	152	-,91746	3,50801	,0000000	1,00000000
Zscore: ChoiceSet1	152	-,10911	10,94746	,0000000	1,00000000
Zscore: ChoiceSet2	152	-7,02424	,14143	,0000000	1,00000000
Zscore: ChoiceSet3	152	-1,77400	1,99731	,0000000	1,00000000
Zscore: ChoiceSet4	152	-,35205	3,10027	,0000000	1,00000000
Zscore: ChoiceSet5	152	-,37388	3,41478	,0000000	1,00000000
Zscore: ChoiceSet6	152	-1,35206	2,16098	,0000000	1,00000000
Zscore: ChoiceSet7	152	-,40273	3,42320	,0000000	1,00000000
Zscore: ChoiceSet8	152	-,36408	3,45244	,0000000	1,00000000
Zscore: ChoiceSet9	152	-3,88587	3,88587	,0000000	1,00000000
Zscore: ChoiceSet10	152	-,84067	1,76712	,0000000	1,00000000
Zscore: ChoiceSet11	152	-,27756	4,16343	,0000000	1,00000000
Zscore: ChoiceSet12	152	-2,09429	2,26642	,0000000	1,00000000
Zscore: ChoiceSet13	152	-1,26117	1,80600	,0000000	1,00000000
Zscore: ChoiceSet14	152	-,89494	2,34388	,0000000	1,00000000
Zscore: ChoiceSet15	152	-,78868	1,41095	,0000000	1,00000000
Zscore: ChoiceSet16	152	-1,54932	1,65472	,0000000	1,00000000
Zscore: ChoiceSet17	152	-1,09945	1,17424	,0000000	1,00000000
Zscore: ChoiceSet18	152	-,43177	3,11577	,0000000	1,00000000
Zscore: Gender	152	-,69783	1,42358	,0000000	1,00000000
Zscore: OS	152	-,83243	5,31829	,0000000	1,00000000
Zscore: NmbrDevices	152	-,88318	2,44516	,0000000	1,00000000
Zscore: Education	152	-1,98269	2,95779	,0000000	1,00000000
Zscore: Profession	152	-1,89345	1,65969	,0000000	1,00000000
Zscore: Income	152	-,74044	2,07322	,0000000	1,00000000
Valid N (listwise)	152				

Report

NmbrDevices

Nationality	Mean	N	Std. Deviation
Dutch	1,93	101	,919
Other European	1,62	13	,961
North America	1,50	4	,577
South America	1,67	12	,778
Asian	1,17	12	,389
Total	1,80	142	,893

Report

Income

OS	Mean	N	Std. Deviation
iOS	2,38	50	1,839
Android	2,34	73	1,773
Windows	1,50	6	,837
iOS and Android	3,40	10	2,366
iOS and Windows	2,00	2	1,414
Android and Windows	2,00	1	.
Total	2,39	142	1,813

VI Analysis in Latent Gold 5.1

A: General Goodness of Fit Statistics

		LL	BIC(LL)	AIC(LL)	AIC3(LL)	df	R^2
Model1	1-Class Choice	-1599,0574	3252,6290	3220,1149	3231,1149	131	0,3916
Model2	2-Class Choice	-1399,1574	2951,9454	2860,3148	2891,3148	111	0,4855
Model3	3-Class Choice	-1292,9220	2838,5912	2687,8440	2738,8440	91	0,5634
Model4	4-Class Choice	-1235,8233	2823,5103	2613,6466	2684,6466	71	0,5865

B: Aggregate 1-class Model

Number of cases	142	
Number of replications	2556	
Number of parameters (Npar)	11	
Random Seed	524717	
Best Start Seed	524717	
Chi-squared Statistics		
Degrees of freedom (df)	131	p-value
L-squared (L^2)	3145,5519	1,4e-567
X-squared	2,522324305e+014	0,0e-2147483647
Cressie-Read	7217580623,7506	4,4e-1567277193
BIC (based on L^2)	2496,3386	
AIC (based on L^2)	2883,5519	
AIC3 (based on L^2)	2752,5519	
CAIC (based on L^2)	2365,3386	
SABIC (based on L^2)	2910,8315	
Dissimilarity Index	0,9994	
Log-likelihood Statistics		
Log-likelihood (LL)	-1599,0574	
Log-prior	-1,4508	
Log-posterior	-1600,5082	
BIC (based on LL)	3252,6290	
AIC (based on LL)	3220,1149	
AIC3 (based on LL)	3231,1149	
CAIC (based on LL)	3263,6290	
SABIC (based on LL)	3217,8242	
Classification Statistics	Classes	
Classification errors	0,0000	
Reduction of errors (Lambda)	1,0000	
Entropy R-squared	1,0000	
Standard R-squared	1,0000	
Classification log-likelihood	-1599,0574	
Entropy	0,0000	
CLC	3198,1149	
AWE	3340,1431	
ICL-BIC	3252,6290	
Classification Table	Modal	
Latent	Class1	Total
Class1	142,0000	142,0000
Total	142,0000	142,0000
Classification Table	Proportional	
Latent	Class1	Total
Class1	142,0000	142,0000
Total	142,0000	142,0000

Model for Choices					
	Class1		Overall		
R²	0,3916		0,3916		
R²(0)	0,4732		0,4732		
Attributes	**Class1**	**z-value**	**Wald**	**p-value**	**Mean**
label					
Label Acknowledged	0,2757	5,6218	35,3558	2,1e-8	0,2757
Label NOT Acknowledged	0,0880	1,2685			0,0880
No Label	-0,3638	-4,5375			-0,3638
rating					
5 Stars	1,4278	26,6130	752,4732	4,0e-164	1,4278
3 Stars	0,1438	1,9176			0,1438
1 Star	-1,5716	-18,6486			-1,5716
reviews					
1 Million Reviews	-0,0761	-1,4138	15,8111	0,00037	-0,0761
500000 Reviews	0,2162	3,8937			0,2162
10000 Reviews	-0,1401	-2,6646			-0,1401
ranking					
Top 3	0,2662	5,0757	65,9607	4,8e-15	0,2662
Top 10	0,2508	4,7179			0,2508
Top 50	-0,5169	-8,1162			-0,5169
price					
0.99	0,9664	16,5734	333,4675	3,9e-73	0,9664
2.99	0,2187	3,4979			0,2187
4.99	-1,1851	-15,8768			-1,1851
none					
	-1,0915	-14,7042	216,2148	6,1e-49	-1,0915

Importance of Factors

	Class1
Maximum	
label	0,6395
rating	2,9993
reviews	0,3563
ranking	0,7831
price	2,1516
none	0,0000
Relative	
label	0,0923
rating	0,4328
reviews	0,0514
ranking	0,1130
price	0,3105
none	0,0000

Prediction Statistics

Choice					
Error Type	Baseline(0)	Baseline	Model	$R^2(0)$	R^2
Squared Error	0,6667	0,5772	0,3512	0,4732	0,3916
Minus Log-likelihood	1,0986	0,9433	0,6256	0,4305	0,3368
Absolute Error	1,3333	1,1575	0,7050	0,4713	0,3909
Prediction Error	0,6667	0,4897	0,2539	0,6191	0,4815
Prediction Table	Estimated				
Observed	1	2	3	Total	
1	1101,0	235,0	0,0	1336,0	
2	147,0	806,0	0,0	953,0	
3	172,0	95,0	0,0	267,0	
Total	1420,0	1136,0	0,0	2556,0	

The prediction table shows that this 1-class (aggregate) model correctly predicts 1101 of the 1336 alternative A responses, and 806 of the 953 alternative B responses. Overall, only 1907 (1101 + 806) of the total 2556 observed choices are predicted correctly.

$$Prediction\ rate = \frac{1101 + 806}{2556} = 74.61\%;\ Prediction\ error = 25.39\%$$

Profile

	Class1
Class Size	1,0000
Attributes	
label	
Label Acknowledged	0,4244
Label NOT Acknowledged	0,3517
No Label	0,2239
rating	
5 Stars	0,7537
3 Stars	0,2087
1 Star	0,0376
reviews	
1 Million Reviews	0,3051
500000 Reviews	0,4087
10000 Reviews	0,2862
ranking	
Top 3	0,4095
Top 10	0,4033
Top 50	0,1872
price	
0.99	0,6290
2.99	0,2978
4.99	0,0732
none	
1	1,0000
Mean	1,0000
Covariates	
Gender	
Male	0,6479
Female	0,3521
Mean	1,3521
Age	
18-24	0,5352
25-34	0,3239
35-49	0,0563
50+	0,0845
Mean	1,6901
Nationality	
Dutch	0,7113
Other European	0,0915
North America	0,0282
South America	0,0845
Asian	0,0845
Mean	1,7394
Education	
1 - 1	0,3521
2 - 2	0,0282
3 - 3	0,4930
4 - 6	0,1268
Mean	3,4507
Profession	
Part-time Employee	0,0986
Full-time Employee	0,2887
Unemployed	0,0211
Student	0,5845
No answer	0,0070
Mean	3,1127

Income	
Less than €1500	0,4789
€1500 - €2999	0,1831
€3000 - €5999	0,1479
€6000 - €8999	0,0211
No answer	0,1690
Mean	2,3873
OS	
1 - 1	0,3521
2 - 2	0,5141
3 - 6	0,1338
Mean	1,9014
Nmbrdevice	
one	0,4507
two	0,3592
three	0,1268
morethan3	0,0634
Mean	1,8028

C: 4-Class Model

Prediction Statistics

Choice					
Error Type	Baseline(0)	Baseline	Model	R²(0)	R²
Squared Error	0,6667	0,5769	0,2386	0,6422	0,5865
Minus Log-likelihood	1,0986	0,9429	0,4226	0,6153	0,5518
Absolute Error	1,3333	1,1548	0,4946	0,6290	0,5717
Prediction Error	0,6667	0,4809	0,1624	0,7565	0,6624
Prediction Table	Estimated				
Observed	1	2	3	Total	
1	1213,0	102,0	21,0	1336,0	
2	111,0	819,0	23,0	953,0	
3	96,0	62,0	109,0	267,0	
Total	1420,0	983,0	153,0	2556,0	

$$Prediction\ rate = \frac{1213 + 819}{2556} = 83.76\%;\ Prediction\ error = 16.24\%$$

Wald p-values in model

Model for Choices										
	Class1	Class2	Class3	Class4	Overall					
R²	0,6692	0,4276	0,6509	0,3197	0,5865					
R²(0)	0,7330	0,4351	0,7590	0,4888	0,6422					
Attributes	**Class1**	**Class2**	**Class3**	**Class4**	**Wald**	**p-value**	**Wald(=)**	**p-value**	**Mean**	**Std.Dev.**
label										
Label Acknowledged	0,0841	0,5205	0,6222	0,8888	65,1959	4,4e-11	18,5767	0,0049	0,3590	0,2818
Label NOT Acknowledged	-0,2942	0,1236	2,2441	0,2853					0,2646	0,8727
No Label	0,2101	-0,6441	-2,8663	-1,1741					-0,6236	1,0675
rating										
5 Stars	2,6025	1,6492	2,6310	0,5929	443,5541	8,9e-91	95,3398	2,3e-18	2,1656	0,6481
3 Stars	-0,0507	0,3649	2,0154	0,5274					0,4344	0,7084
1 Star	-2,5518	-2,0141	-4,6464	-1,1203					-2,6000	0,9688
reviews										
1 Million Reviews	-0,1378	0,1743	-0,4505	0,2207	16,8276	0,032	11,3079	0,079	-0,0685	0,2203
500000 Reviews	0,1412	-0,0100	1,3258	0,4055					0,3065	0,4496
10000 Reviews	-0,0033	-0,1643	-0,8752	-0,6262					-0,2380	0,3229
ranking										
Top 3	0,1138	0,4203	2,3220	0,4447	39,0846	4,7e-6	8,9029	0,18	0,5689	0,7631
Top 10	0,0294	0,2545	1,6637	0,2228					0,3611	0,5658
Top 50	-0,1432	-0,6749	-3,9857	-0,6675					-0,9299	1,3288
price										
0.99	0,9874	1,2972	4,5785	0,5180	197,1543	2,5e-38	22,8958	0,00083	1,5854	1,2969
2.99	-0,1722	0,1797	2,0128	0,3637					0,3100	0,7518
4.99	-0,8151	-1,4769	-6,5913	-0,8817					-1,8955	2,0275
none										
	-1,1751	0,4404	-7,6443	-8,3667	57,3871	1,0e-11	57,0198	2,5e-12	-2,3761	3,2208

Z-values in model

Model for Choices										
	Class1		Class2		Class3		Class4		Overall	
R²	0,6692		0,4276		0,6509		0,3197		0,5865	
R²(0)	0,7330		0,4351		0,7590		0,4888		0,6422	
Attributes	**Class1**	**z-value**	**Class2**	**z-value**	**Class3**	**z-value**	**Class4**	**z-value**	**Mean**	**Std.Dev.**
label										
Label Acknowledged	0,0841	0,5068	0,5205	4,5515	0,6222	2,6915	0,8888	5,3767	0,3590	0,2818
Label NOT Acknowledged	-0,2942	-1,7281	0,1236	0,9473	2,2441	1,3768	0,2853	1,0188	0,2646	0,8727
No Label	0,2101	0,8738	-0,6441	-4,1289	-2,8663	-1,5988	-1,1741	-3,4650	-0,6236	1,0675
rating										
5 Stars	2,6025	15,0545	1,6492	13,3470	2,6310	1,9470	0,5929	3,5324	2,1656	0,6481
3 Stars	-0,0507	-0,2528	0,3649	2,3299	2,0154	2,3690	0,5274	1,8263	0,4344	0,7084
1 Star	-2,5518	-10,9348	-2,0141	-10,9567	-4,6464	-2,1610	-1,1203	-3,1307	-2,6000	0,9688
reviews										
1 Million Reviews	-0,1378	-0,9470	0,1743	1,5466	-0,4505	-1,5932	0,2207	1,3741	-0,0685	0,2203
500000 Reviews	0,1412	1,0767	-0,0100	-0,0761	1,3258	1,4263	0,4055	1,9828	0,3065	0,4496
10000 Reviews	-0,0033	-0,0230	-0,1643	-1,3589	-0,8752	-1,1768	-0,6262	-3,0868	-0,2380	0,3229
ranking										
Top 3	0,1138	0,8433	0,4203	3,6633	2,3220	1,8838	0,4447	2,5481	0,5689	0,7631
Top 10	0,0294	0,2148	0,2545	2,2809	1,6637	1,9530	0,2228	1,3070	0,3611	0,5658
Top 50	-0,1432	-0,8409	-0,6749	-4,7565	-3,9857	-1,9444	-0,6675	-2,7920	-0,9299	1,3288
price										
0.99	0,9874	6,4444	1,2972	9,9721	4,5785	2,7647	0,5180	2,9676	1,5854	1,2969
2.99	-0,1722	-0,9289	0,1797	1,4633	2,0128	2,1651	0,3637	1,4738	0,3100	0,7518
4.99	-0,8151	-4,4691	-1,4769	-9,8410	-6,5913	-2,5841	-0,8817	-2,7888	-1,8955	2,0275
none										
	-1,1751	-5,9064	0,4404	3,3051	-7,6443	-2,9191	-8,3667	-1,3396	-2,3761	3,2208

Importance

	Class1	Class2	Class3	Class4
Maximum				
label	0,5043	1,1645	5,1104	2,0628
rating	5,1543	3,6633	7,2774	1,7133
reviews	0,2790	0,3386	2,2010	1,0318
ranking	0,2571	1,0952	6,3078	1,1121
price	1,8025	2,7742	11,1697	1,3996
none	0,0000	0,0000	0,0000	0,0000
Relative				
label	0,0631	0,1289	0,1594	0,2818
rating	0,6445	0,4054	0,2269	0,2341
reviews	0,0349	0,0375	0,0686	0,1410
ranking	0,0321	0,1212	0,1967	0,1519
price	0,2254	0,3070	0,3483	0,1912
none	0,0000	0,0000	0,0000	0,0000

D: Refined 4-class model

Goodness of Fit

		LL	BIC(LL)	AIC(LL)	AIC3(LL)	CAIC(LL)	L²	df	R²
Model1	1-Class Choice	-1599,0574	3252,6290	3220,1149	3231,1149	3263,6290	3145,5519	131	0,3916
Model2	2-Class Choice	-1399,1574	2951,9454	2860,3148	2891,3148	2982,9454	2745,7519	111	0,4855
Model3	3-Class Choice	-1292,9220	2838,5912	2687,8440	2738,8440	2889,5912	2533,2811	91	0,5634
Model4	4-Class Choice	-1235,8233	2823,5103	2613,6466	2684,6466	2894,5103	2419,0836	71	0,5865
Refined	4-Class Choice	-1240,8818	2793,9808	2607,7637	2670,7637	2856,9808	2429,2007	79	0,5857

Prediction Statistics

Choice					
Error Type	Baseline(0)	Baseline	Model	R²(0)	R²
Squared Error	0,6667	0,5769	0,2390	0,6415	0,5857
Minus Log-likelihood	1,0986	0,9430	0,4235	0,6146	0,5509
Absolute Error	1,3333	1,1554	0,4966	0,6275	0,5702
Prediction Error	0,6667	0,4829	0,1635	0,7547	0,6614
Prediction Table	Estimated				
Observed	1	2	3	Total	
1	1210,0	104,0	22,0	1336,0	
2	112,0	820,0	21,0	953,0	
3	96,0	63,0	108,0	267,0	
Total	1418,0	987,0	151,0	2556,0	

Prediction rate = 83.65%; Prediction error = 16.35%

Wald p-values in refined model

Model for Choices										
	Class1	Class2	Class3	Class4	Overall					
R²	0,6649	0,4268	0,6439	0,3214	0,5857					
R²(0)	0,7296	0,4337	0,7537	0,4901	0,6415					
Attributes	**Class1**	**Class2**	**Class3**	**Class4**	**Wald**	**p-value**	**Wald(=)**	**p-value**	**Mean**	**Std.Dev.**
label										
Label Acknowledged	0,1710	0,4998	0,4185	0,8878	76,6279	2,3e-13	21,2310	0,0017	0,3624	0,2195
Label NOT Acknowledged	-0,3952	0,1213	0,2823	0,2820					-0,0900	0,3021
No Label	0,2243	-0,6211	-0,7008	-1,1698					-0,2724	0,5035
rating										
5 Stars	2,6226	1,7110	0,9745	0,6037	502,1471	2,4e-103	106,1893	1,3e-20	1,9397	0,7387
3 Stars	-0,0644	0,3243	1,0758	0,5297					0,2707	0,4025
1 Star	-2,5581	-2,0353	-2,0503	-1,1334					-2,2104	0,4187
reviews										
1 Million Reviews	0,0000	0,0000	0,0000	0,2091	9,5033	0,0086	0,0000	.	0,0188	0,0598
500000 Reviews	0,0000	0,0000	0,0000	0,4125					0,0371	0,1180
10000 Reviews	-0,0000	-0,0000	-0,0000	-0,6216					-0,0559	0,1779
ranking										
Top 3	0,0000	0,4201	0,9037	0,4516	57,2490	1,6e-10	5,1819	0,27	0,2940	0,3259
Top 10	0,0000	0,2669	0,6016	0,2202					0,1851	0,2133
Top 50	-0,0000	-0,6870	-1,5052	-0,6718					-0,4791	0,5385
price										
0.99	1,0210	1,3340	2,5673	0,5129	287,2096	2,2e-57	46,0424	2,9e-8	1,2996	0,5859
2.99	-0,2604	0,1640	1,0315	0,3642					0,1106	0,4528
4.99	-0,7606	-1,4979	-3,5987	-0,8771					-1,4103	0,9881
none										
	-1,1817	0,5172	-5,1483	-8,3770	87,0287	5,6e-18	86,2041	1,4e-18	-1,9877	2,6886

Z-Values in refined model

Model for Choices										
	Class1		Class2		Class3		Class4		Overall	
R²	0,6649		0,4268		0,6439		0,3214		0,5857	
R²(0)	0,7296		0,4337		0,7537		0,4901		0,6415	
Attributes	**Class1**	**z-value**	**Class2**	**z-value**	**Class3**	**z-value**	**Class4**	**z-value**	**Mean**	**Std.Dev.**
label										
Label Acknowledged	0,1710	1,2822	0,4998	4,7136	0,4185	2,7495	0,8878	5,4171	0,3624	0,2195
Label NOT Acknowledged	-0,3952	-3,0417	0,1213	0,9613	0,2823	0,8614	0,2820	1,0139	-0,0900	0,3021
No Label	0,2243	1,1517	-0,6211	-4,1484	-0,7008	-2,1223	-1,1698	-3,4564	-0,2724	0,5035
rating										
5 Stars	2,6226	16,1311	1,7110	13,9185	0,9745	4,0838	0,6037	3,6498	1,9397	0,7387
3 Stars	-0,0644	-0,3419	0,3243	2,0981	1,0758	2,8955	0,5297	1,8458	0,2707	0,4025
1 Star	-2,5581	-11,0941	-2,0353	-10,9254	-2,0503	-4,2101	-1,1334	-3,2124	-2,2104	0,4187
reviews										
1 Million Reviews	0,0000	.	0,0000	.	0,0000	.	0,2091	1,3050	0,0188	0,0598
500000 Reviews	0,0000	.	0,0000	.	0,0000	.	0,4125	2,0151	0,0371	0,1180
10000 Reviews	-0,0000	.	-0,0000	.	-0,0000	.	-0,6216	-3,0791	-0,0559	0,1779
ranking										
Top 3	0,0000	.	0,4201	3,6529	0,9037	3,3968	0,4516	2,6047	0,2940	0,3259
Top 10	0,0000	.	0,2669	2,5806	0,6016	2,9653	0,2202	1,2964	0,1851	0,2133
Top 50	-0,0000	.	-0,6870	-5,3703	-1,5052	-4,3586	-0,6718	-2,8288	-0,4791	0,5385
price										
0.99	1,0210	8,1264	1,3340	10,2477	2,5673	7,4430	0,5129	2,9571	1,2996	0,5859
2.99	-0,2604	-1,8132	0,1640	1,3348	1,0315	3,1582	0,3642	1,5136	0,1106	0,4528
4.99	-0,7606	-5,8320	-1,4979	-9,8919	-3,5987	-6,1322	-0,8771	-2,8520	-1,4103	0,9881
none										
	-1,1817	-6,4067	0,5172	3,8343	-5,1483	-4,9589	-8,3770	-1,3413	-1,9877	2,6886

Importance of factors

	Class1	Class2	Class3	Class4
Maximum				
label	0,6195	1,1210	1,1193	2,0576
rating	5,1807	3,7464	3,1260	1,7371
reviews	0,0000	0,0000	0,0000	1,0341
ranking	0,0000	1,1072	2,4089	1,1234
price	1,7816	2,8319	6,1660	1,3900
none	0,0000	0,0000	0,0000	0,0000
Relative				
label	0,0817	0,1273	0,0873	0,2802
rating	0,6833	0,4254	0,2438	0,2366
reviews	0,0000	0,0000	0,0000	0,1408
ranking	0,0000	0,1257	0,1879	0,1530
price	0,2350	0,3216	0,4810	0,1893
none	0,0000	0,0000	0,0000	0,0000

Wald p-values of covariates in Refined Model

Model for Classes						
Intercept	**Class1**	**Class2**	**Class3**	**Class4**	**Wald**	**p-value**
	2,0028	0,6519	-0,1487	-2,5061	3,2364	0,36
Covariates	**Class1**	**Class2**	**Class3**	**Class4**	**Wald**	**p-value**
Gender						
	-1,0759	-0,1631	0,9518	0,2872	12,4518	0,0060
Age						
	-0,1236	0,0443	0,4116	-0,3323	2,1492	0,54
Nationality						
	-0,1484	0,2400	-0,1982	0,1066	5,3204	0,15
Education						
	-0,0800	-0,1856	-0,6666	0,9323	9,8918	0,020
Profession						
	0,0603	0,1166	0,1657	-0,3426	1,4933	0,68
Income						
	0,0835	0,0011	0,1574	-0,2420	1,9600	0,58
OS						
	0,1970	-0,2986	0,2116	-0,1100	2,9032	0,41
Nmbrdevice						
	0,3136	0,2526	-0,4868	-0,0793	3,6340	0,30

Profile

	Class1	Class2	Class3	Class4
Overall	0,4856	0,2692	0,1553	0,0900
Attributes				
label				
Label Acknowledged	0,4176	0,3019	0,1593	0,1213
Label NOT Acknowledged	0,3653	0,3186	0,2141	0,1020
No Label	0,7264	0,1623	0,0858	0,0255
rating				
5 Stars	0,5810	0,2715	0,0926	0,0549
3 Stars	0,1516	0,2601	0,3927	0,1955
1 Star	0,1371	0,2689	0,1886	0,4055
reviews				
1 Million Reviews	0,4801	0,2661	0,1535	0,1003
500000 Reviews	0,4695	0,2602	0,1501	0,1202
10000 Reviews	0,5089	0,2821	0,1627	0,0463
ranking				
Top 3	0,3926	0,2984	0,2059	0,1030
Top 10	0,4449	0,2900	0,1725	0,0926
Top 50	0,7227	0,1815	0,0341	0,0617
price				
0.99	0,4780	0,2798	0,1816	0,0607
2.99	0,4268	0,2793	0,1257	0,1681
4.99	0,7157	0,1466	0,0034	0,1344
none				
1	0,4856	0,2692	0,1553	0,0900
Covariates				
Gender				
Male	0,6192	0,2325	0,0749	0,0734
Female	0,2404	0,3366	0,3027	0,1203
Age				
18-24	0,5244	0,2763	0,1457	0,0536
25-34	0,3644	0,2965	0,1947	0,1444
35-49	0,8188	0,1797	0,0000	0,0015
50+	0,4835	0,1785	0,1682	0,1697
Nationality				
Dutch	0,5477	0,2071	0,1593	0,0859
Other European	0,4386	0,3325	0,2270	0,0019
North America	0,4976	0,2497	0,0000	0,2528
South America	0,2472	0,4176	0,1671	0,1681
Asian	0,2482	0,5810	0,0840	0,0868
Education				
1 - 1	0,4804	0,2576	0,2615	0,0005
2 - 2	0,4985	0,2502	0,2513	0,0000
3 - 3	0,4835	0,2917	0,1008	0,1241
4 - 6	0,5059	0,2185	0,0514	0,2242
Profession				
Part-time Employee	0,4829	0,2277	0,0722	0,2172
Full-time Employee	0,5558	0,1865	0,1699	0,0879
Unemployed	0,6627	0,0000	0,0000	0,3373
Student	0,4388	0,3300	0,1696	0,0616
No answer	0,9994	0,0003	0,0000	0,0003
Income				
Less than €1500	0,4701	0,3207	0,1490	0,0602
€1500 - €2999	0,4365	0,1605	0,1546	0,2484
€3000 - €5999	0,4797	0,2822	0,1849	0,0532
€6000 - €8999	0,8519	0,1447	0,0000	0,0034
No answer	0,5422	0,2453	0,1673	0,0451

OS				
1 - 1	0,4758	0,3232	0,1010	0,0999
2 - 2	0,4053	0,2830	0,2052	0,1065
3 - 6	0,8208	0,0733	0,1058	0,0001
Nmbrdevice				
one	0,3424	0,3045	0,2015	0,1516
two	0,6267	0,2356	0,1186	0,0191
three	0,4563	0,2594	0,1721	0,1121
morethan3	0,7654	0,2268	0,0000	0,0078

E: Alternative 3-class model

Prediction rate = 82.63%; Prediction error = 17.37%

Statistics of a 3-class model

Model for Choices								
	Class1		Class2		Class3		Overall	
R²	0,6676		0,3816		0,4493		0,5634	
R²(0)	0,7323		0,5399		0,4551		0,6221	
Attributes	Class1	z-value	Class2	z-value	Class3	z-value	Wald	p-value
label								
Label Acknowledged	0,1331	0,8406	0,4495	5,1635	0,5499	4,2898	52,1919	1,7e-9
Label NOT Acknowledged	-0,2854	-1,6829	0,2496	1,5081	0,0748	0,5291		
No Label	0,1523	0,6569	-0,6991	-3,6946	-0,6246	-3,6600		
rating								
5 Stars	2,5607	16,0857	0,5366	4,8976	1,8173	12,9831	475,1059	1,9e-99
3 Stars	-0,0435	-0,2239	0,4670	2,9906	0,4379	2,5189		
1 Star	-2,5173	-11,1224	-1,0036	-5,0930	-2,2552	-10,4821		
reviews								
1 Million Reviews	-0,1323	-0,9395	0,0717	0,6999	0,1964	1,5964	12,0132	0,062
500000 Reviews	0,1587	1,1987	0,2273	1,9697	-0,0256	-0,1774		
10000 Reviews	-0,0264	-0,1874	-0,2990	-2,6382	-0,1708	-1,2933		
ranking								
Top 3	0,1147	0,8979	0,4646	4,0824	0,3983	3,2846	46,2500	2,6e-8
Top 10	0,0518	0,3901	0,2762	2,6672	0,2726	2,2439		
Top 50	-0,1665	-1,0438	-0,7409	-4,8222	-0,6709	-4,3822		
price								
0.99	0,9741	6,6587	1,2532	10,2658	1,3002	8,9606	271,8046	8,9e-56
2.99	-0,1440	-0,8092	0,3687	2,9219	0,1921	1,4369		
4.99	-0,8301	-4,5721	-1,6220	-8,6629	-1,4924	-8,9665		
none								
	-1,2396	-6,0013	-3,9288	-8,7780	0,6546	4,2955	144,3792	4,3e-31

Importance of factors

	Class1	Class2	Class3
Maximum			
label	0,4378	1,1485	1,1745
rating	5,0780	1,5402	4,0725
reviews	0,2910	0,5263	0,3672
ranking	0,2812	1,2055	1,0692
price	1,8042	2,8752	2,7926
none	0,0000	0,0000	0,0000
Relative			
label	0,0555	0,1574	0,1239
rating	0,6434	0,2111	0,4298
reviews	0,0369	0,0721	0,0388
ranking	0,0356	0,1652	0,1128
price	0,2286	0,3941	0,2947
none	0,0000	0,0000	0,0000

F: Adapted profiles of the 4 segments

	Class1	Class2	Class3	Class4
Class Size	0,4809	0,2678	0,1544	0,0969
Attributes				
label				
Label Acknowledged	0,3590	0,5087	0,1645	0,5957
Label NOT Acknowledged	0,2459	0,3315	0,8305	0,3040
No Label	0,3951	0,1598	0,0050	0,1002
rating				
5 Stars	0,9280	0,7832	0,6487	0,4799
3 Stars	0,0666	0,2017	0,3508	0,4000
1 Star	0,0054	0,0151	0,0004	0,1201
reviews				
1 Million Reviews	0,2880	0,4002	0,1322	0,3806
500000 Reviews	0,3828	0,3317	0,7812	0,4161
10000 Reviews	0,3292	0,2680	0,0865	0,2033
ranking				
Top 3	0,3719	0,4390	0,6578	0,4811
Top 10	0,3437	0,3989	0,3410	0,3528
Top 50	0,2844	0,1621	0,0012	0,1661
price				
0.99	0,6753	0,7290	0,9287	0,4911
2.99	0,2147	0,2288	0,0713	0,3630
4.99	0,1100	0,0422	0,0000	0,1459
none				
1	1,0000	1,0000	1,0000	1,0000
Mean	1,0000	1,0000	1,0000	1,0000
Covariates				
Gender				
Male	0,8238	0,5834	0,3112	0,4824
Female	0,1762	0,4166	0,6888	0,5176
Mean	1,1762	1,4166	1,6888	1,5176
Age				
18-24	0,5741	0,5389	0,5034	0,3800
25-34	0,2473	0,3635	0,4045	0,4705
35-49	0,1787	0,0976	0,0921	0,1495
Mean	1,6046	1,5587	1,5887	1,7694
Nationality				
Dutch	0,7985	0,5774	0,7267	0,6214
Other European	0,0847	0,1071	0,1359	0,0112
North America	0,0292	0,0263	0,0000	0,0735
South America	0,0435	0,1321	0,0916	0,1469
Asian	0,0441	0,1571	0,0458	0,1470
Mean	1,4500	2,1844	1,5939	2,1867
Education				
High School	0,3698	0,3583	0,6408	0,0764
Bachelor's Degree	0,4978	0,5382	0,3219	0,6164
Master's Degree	0,0868	0,0804	0,0374	0,1600
Doctor's Degree	0,0455	0,0231	0,0000	0,0003
Other	0,0000	0,0000	0,0000	0,1470
Mean	3,4382	3,4098	2,7559	4,4486
Profession				
Part-time Employee	0,1002	0,0821	0,0465	0,2202
Full-time Employee	0,3348	0,2046	0,3129	0,2533
Unemployed	0,0292	0,0000	0,0000	0,0735
Student	0,5212	0,7133	0,6407	0,4529
No answer	0,0146	0,0000	0,0000	0,0000

Income				
Less than €1500	0,4560	0,5931	0,4588	0,3070
€1500 - €2999	0,1715	0,1080	0,1841	0,4498
€3000 - €5999	0,1820	0,1710	0,1739	0,0902
No answer	0,1905	0,1279	0,1833	0,1530
Mean	2,4880	2,0893	2,4482	2,3952
OS				
1 - 1	0,3392	0,4150	0,2226	0,4493
2 - 2	0,4322	0,5504	0,6855	0,5502
3 - 6	0,2286	0,0347	0,0919	0,0005
Mean	2,0927	1,6475	1,9609	1,5520
Nmbrdevice				
one	0,3246	0,4837	0,5862	0,7763
two	0,4558	0,3361	0,2754	0,0711
three	0,1194	0,1264	0,1384	0,1463
morethan3	0,1002	0,0539	0,0000	0,0062
Mean	1,9951	1,7505	1,5523	1,3824