

## IDENTIFICATION OF QUALITY CHARACTERISTICS FOR MOBILE APPLICATIONS

Vitor Carneiro Maia

Master's dissertation presented to the Graduate Program in Systems and Computing Engineering, COPPE, Federal University of Rio de Janeiro, as part of the requirements to obtain the title of Master in Systems and Computing Engineering.

Advisors: Ana Regina Cavalcanti da Rocha  
Taisa Guidini Gonçalves

Rio de Janeiro  
December, 2019

## IDENTIFICATION OF QUALITY CHARACTERISTICS FOR MOBILE APPLICATIONS

Vitor Carneiro Maia

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Engenharia de Sistemas e Computação, COPPE, da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Mestre em Engenharia de Sistemas e Computação.

Orientadores: Ana Regina Cavalcanti da Rocha  
Taisa Guidini Gonçalves

Rio de Janeiro  
Dezembro de 2019

IDENTIFICATION OF QUALITY CHARACTERISTICS FOR MOBILE  
APPLICATIONS

Vitor Carneiro Maia

DISSERTAÇÃO SUBMETIDA AO CORPO DOCENTE DO INSTITUTO  
ALBERTO LUIZ COIMBRA DE POS-GRADUAÇÃO E PESQUISA DE  
ENGENHARIA (COPPE) DA UNIVERSIDADE FEDERAL DO RIO DE  
JANEIRO COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A  
OBTENÇÃO DO GRAU DE MESTRE EM CIÊNCIAS EM ENGENHARIA DE  
SISTEMAS E COMPUTAÇÃO.

Examinada por:

---

Profa. Ana Regina Cavalcanti da Rocha

---

Prof. Guilherme Horta Travassos

---

Prof. Gleison dos Santos Souza

RIO DE JANEIRO, RJ – BRASIL

DEZEMBRO DE 2019

Maia, Vitor Carneiro

Identification of Quality Characteristics for Mobile Applications – Rio de Janeiro: UFRJ/COPPE, 2019

X, 109 p.: il.; 29,7 cm.

Orientadores: Ana Regina Cavalcanti da Rocha

Taisa Guidini Gonçalves

Dissertação (mestrado) – UFRJ/COPPE/Programa de Engenharia de Sistemas e Computação, 2019.

Referências Bibliográficas: p. 82 – 93.

1. Engenharia de Software. 2. Qualidade de Software. 3. Aplicações Móveis. I. da Rocha, Ana Regina Cavalcanti et al. II. Universidade Federal do Rio de Janeiro, COPPE, Programa de Engenharia de Sistemas e Computação. III. Título.

Resumo da Dissertação apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Mestre em Ciências (M.Sc.)

## IDENTIFICAÇÃO DE CARACTERÍSTICAS DE QUALIDADE PARA APLICAÇÕES MÓVEIS

Vitor Carneiro Maia

Dezembro/2019

Orientadores: Ana Regina Cavalcanti da Rocha

Taisa Guidini Gonçalves

Programa: Engenharia de Sistemas e Computação

Aplicações móveis possuem um mercado bem estabelecido e tornaram-se populares após a criação do modelo de distribuição por lojas de aplicativos. Estes sistemas são diretamente impactados por variações de contexto, pelo uso de sensores presentes nos dispositivos móveis, pela usabilidade, além de outras particularidades que as diferenciam de outros tipos de software. Lojas de aplicativos possuem diretrizes de qualidade para publicação com recomendações genéricas, mas não abrangem todas as necessidades e formas de uso das aplicações móveis. Modelos de qualidade para avaliação da qualidade em produtos de software são genéricos, apresentam características de qualidade tendo em vista qualquer tipo de aplicação. Porém, estes modelos podem ser particularizados para uso apenas em contextos específicos. Este trabalho propõe um conjunto de características de qualidade específicas para o contexto de aplicações móveis a partir da particularização dos modelos de qualidade definidos em normas internacionais. Este conjunto foi utilizado para estender os modelos de qualidade da ISO/IEC 25010. A identificação foi realizada através da condução de um mapeamento sistemático e de um survey com usuários de aplicações móveis. Com o conjunto de características de qualidade identificado, é também proposto um procedimento de avaliação específico para aplicações móveis, adaptado de um modelo de avaliação de produtos de software já existente. A viabilidade de uso deste modelo foi verificada através da avaliação de uma aplicação móvel bancária, disponível no mercado.

Abstract of Dissertation presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Master of Science (M.Sc.)

## IDENTIFICATION OF QUALITY CHARACTERISTICS FOR MOBILE APPLICATIONS

Vitor Carneiro Maia

December/2019

Advisors: Ana Regina Cavalcanti da Rocha

Taisa Guidini Gonçalves

Department: Systems Engineering and Computer Science

Mobile applications hold a well-established market, and they became popular since the creation of the app stores' distribution model. These systems are directly impacted by context variations, by the usage of mobile device's sensors, by usability and many other particular features that makes mobile applications different from other types of software. App stores provide quality guidelines with generic recommendations for publishing apps, but these recommendations do not cover all mobile application needs and usages. Quality models for quality assessment of software product are general-purpose and made to be applied to any kind of software product. However, these models may be particularized for specific contexts. This paper proposed a set of context-specific quality characteristics for mobile applications, based on the particularization of quality models defined in international standards. This set was used to extend the ISO/IEC 25010 quality models. The identification was performed by conducting a systematic mapping and a survey with mobile application users. One the quality characteristics are identified, an evaluation procedure for mobile application is also proposed. The procedure is adapted from an existing software evaluation model. The feasibility of using this procedure was verified through the evaluation of a mobile banking application, already on the market.

# Index

Index of Figures .....	viii
Index of Tables .....	ix
<b>1 Introduction .....</b>	<b>1</b>
<b>1.1 Context .....</b>	<b>1</b>
<b>1.2 Motivation and Objective .....</b>	<b>2</b>
<b>1.3 Methodology .....</b>	<b>3</b>
<b>1.4 Dissertation's Organization .....</b>	<b>4</b>
<b>2 Software Product Quality .....</b>	<b>5</b>
<b>2.1 Software Quality .....</b>	<b>5</b>
<b>2.2 Historical Quality Models for Software Product .....</b>	<b>5</b>
<b>2.3 International Standards for Product Quality .....</b>	<b>7</b>
<b>2.3.1 International Standard ISO/IEC 25010 .....</b>	<b>8</b>
<b>2.3.2 International Standard ISO/IEC 25051 .....</b>	<b>12</b>
<b>2.4 Software Quality Evaluation and Certification .....</b>	<b>13</b>
<b>2.4.1 National Context .....</b>	<b>13</b>
<b>2.4.2 International Context .....</b>	<b>14</b>
<b>2.4.3 QPS .....</b>	<b>15</b>
<b>2.5 Particularizations of Quality Models .....</b>	<b>17</b>
<b>2.6 Conclusion .....</b>	<b>19</b>
<b>3 Systematic Mapping .....</b>	<b>20</b>
<b>3.1 Research Method .....</b>	<b>20</b>
<b>3.2 Research Goal .....</b>	<b>21</b>
<b>3.3 Planning .....</b>	<b>22</b>
<b>3.3.1 Research Questions .....</b>	<b>22</b>
<b>3.3.2 Search String .....</b>	<b>23</b>
<b>3.3.3 Selection of Search Engines .....</b>	<b>23</b>
<b>3.3.4 Inclusion and Exclusion Criteria .....</b>	<b>23</b>
<b>3.4 Execution .....</b>	<b>24</b>
<b>3.5 Analysis of the Results .....</b>	<b>30</b>
<b>3.5.1 Data Extraction Form .....</b>	<b>30</b>
<b>3.5.2 RQ1: Quality Characteristics from ISO/IEC 25010 .....</b>	<b>30</b>
<b>3.5.3 RQ2: Quality Characteristics not present in ISO/IEC 25010 .....</b>	<b>34</b>
<b>3.5.4 Additional Results .....</b>	<b>36</b>
<b>3.6 Discussion of the Results .....</b>	<b>40</b>
<b>3.7 Threats to Validity .....</b>	<b>41</b>

3.8	Conclusion.....	42
4	Survey.....	43
4.1	Goal and Sub-Characteristics .....	43
4.2	Instrument .....	45
4.2.1	Target Audience .....	48
4.2.2	Pilot Test.....	48
4.3	Execution.....	48
4.4	Data Analysis .....	49
4.5	Discussion.....	52
4.5.1	Favorite App Category .....	52
4.5.2	Characterization.....	54
4.5.3	Answer Time and Answers per day .....	54
4.5.4	Multiple Choice Questions.....	55
4.5.5	Quality Characteristics per App Category .....	56
4.6	Threats to Validity .....	59
4.7	Final List of Essential Sub-characteristics.....	60
4.8	Conclusion.....	63
5	Quality Evaluation of Mobile Applications.....	64
5.1	The QPS process and its appraisal method.....	64
5.2	Appraisal of Mobile Applications .....	68
5.3	Appraisal Planning and Execution for a Mobile Applications.....	73
5.3.1	Appraisal Plan .....	73
5.3.2	Execution of the Appraisal .....	73
5.3.3	Appraisal Team’s Feedback concerning the Appraisal .....	76
5.4	Conclusion.....	77
6	Conclusion.....	78
6.1	Obtained Results .....	78
6.2	Limitations .....	79
6.3	Bibliographic Production .....	80
6.4	Future Work .....	80
	Bibliographical References.....	82
	Annex A – Extraction Form .....	94
	Annex B – Appraisal Instrument.....	95
	Annex C - Appraisal Plan.....	99
	Annex D – Final Report (Filled).....	101
	Annex E - Revised Instrument .....	105



# Index of Figures

Figure 1.1 – Methodology .....	3
Figure 2.1 – McCall Quality Model (McCall, 1977).....	6
Figure 2.2 – Boehm Quality Model (Boehm, 1978) .....	7
Figure 2.3 – ISO/IEC 9126 External and Internal Quality Model (ISO/IEC, 2001).....	8
Figure 2.4 – ISO/IEC 25010 Product Quality Model (ISO/IEC, 2011) .....	9
Figure 2.5 – ISO/IEC 25010 Quality in Use Model (ISO/IEC, 2011) .....	11
Figure 2.6 – Overview of QPS Model and the relation with international standards.....	16
Figure 3.1 – Summary of the Systematic Review Process (Biolchini et al., 2005).....	20
Figure 3.2 – Percentage of use of each software quality standard. ....	39
Figure 3.3 – Frequency of publications over the years .....	40
Figure 4.1 – Example of VAS question .....	47
Figure 4.2 – Gender of the participants.....	49
Figure 4.3 – Education Level of the participants .....	50
Figure 4.4 – Age of the participants.....	50
Figure 4.5 – App Categories selected by the participants .....	51
Figure 4.6 – Average of the quality characteristics per app category .....	53
Figure 4.7 – Quantity of answers per day .....	55
Figure 4.8 – Overall Average of the quality characteristics .....	58
Figure 4.9 – Customized versions of ISO/IEC 25010 quality models .....	63
Figure 5.1 –Initial Diagnosis Activities and Tasks .....	65
Figure 5.2 – Final Assessment Activities and Tasks.....	66
Figure 5.3 – Procedure for the initial characterization of products already on the market .....	66
Figure 5.4 – Activities and Tasks used in the Final Appraisal (apps in an app store).....	69

# Index of Tables

Table 2.1 – Characteristics and Sub-characteristics in the Product Quality Model .....	9
Table 2.2 – Characteristics and Sub-characteristics in the Quality in Use Model .....	11
Table 2.3 – Examples of Product Description Requirements.....	12
Table 2.4 – Examples of User Documentation Requirements .....	13
Table 2.5 – Examples of Software Requirements .....	13
Table 2.6 – Examples of checklist items in MEDE-PROS guide .....	14
Table 3.1 – Research goal, according to GQM paradigm .....	22
Table 3.2 – Research Questions .....	22
Table 3.3 – Exclusion Criteria.....	24
Table 3.4 – Inclusion Criteria.....	24
Table 3.5 – Search Results .....	24
Table 3.6 – Excluded papers and the engines where they were found.....	25
Table 3.7 – Papers included via search string and where they were found.....	25
Table 3.8 – Papers included via Snowballing, a manual search in the search engines .....	26
Table 3.9 – Summary of the included papers .....	26
Table 3.10 – Occurrences of Product Quality model’s characteristics and sub-characteristics. .	32
Table 3.11 – Occurrences of Quality in Use model’s characteristics and sub-characteristics ....	33
Table 3.12 – Product quality characteristics not present in ISO/IEC 25010.....	35
Table 3.13 – Quality in use characteristics not present in ISO/IEC 25010. ....	36
Table 3.14 – Definitions of specific mobile application categories .....	36
Table 3.15 – Definitions of mobile application.....	38
Table 4.1 – Quality attributes, not present in ISO/IEC 25010 quality models.....	44
Table 4.2 – Quality sub-characteristics not considered essential in the context of apps (< 5)....	44
Table 4.3 – Quality sub-characteristics considered essential in the context of apps (>= 8).....	45
Table 4.4 – Quality sub-characteristics added to the survey for further investigation (5 - 7) .....	45
Table 4.5 – Survey questions for the ISO/IEC 25010 quality sub-characteristics .....	46
Table 4.6 – Survey questions for the quality attributes not previewed by ISO/IEC 25010 .....	47
Table 4.7 – Final list of essential characteristics for the context of mobile applications .....	61
Table 4.8 – Dismissed quality characteristics and their ids in Figure 4.9 .....	62
Table 4.9 – Dismissed quality sub-characteristics, and their ids in Figure 4.9 .....	62
Table 5.1 – Characterization Rules .....	67
Table 5.2 – Aggregation rules .....	67
Table 5.3 – Questions for the assessment of product quality sub-characteristics .....	70
Table 5.4 – Questions for the evaluation of quality in use sub-characteristics .....	71
Table 5.5 – Characterization of the questions of the product quality sub-characteristics .....	74

Table 5.6 – Characterization of the questions of quality in use sub-characteristics.....	75
Table 5.7 – Characterization of the product quality characteristics .....	75
Table 5.8 – Characterization of the product quality characteristics .....	76

# 1 Introduction

This dissertation proposes a set of quality characteristics for the context of mobile applications. Given this set of quality characteristics, the dissertation also proposes an adaptation of ISO/IEC 25010 (ISO/IEC, 2011) quality models for the context of mobile applications and an appraisal procedure. Section 1.1 introduces the context in which the dissertation's theme is located. Section 1.2 presents the motivation and objective of this work. Section 1.3 presents the dissertation's methodology. Section 1.4 presents a brief description of each of the following chapters.

## 1.1 Context

Quality Models for evaluating software quality are general-purpose. However, different applications and stakeholders demand context-specific quality requirements. It becomes crucial to identify specific quality requirements for these types of application, as well as stakeholders with whom draw up quality characteristics, sub-characteristics and evaluation procedures.

Mobile applications hold a robust and established market. Mobile applications became very popular since the creation of the app stores' distribution model. Both the quantity of apps and the variety of their functionalities highly increased ever since. The biggest app stores (environments where mobile applications are published) are for Android and iOS systems. Both platforms provide publishing quality guidelines, but they are not broad enough to grant the final users' expectations and app quality.

The Android app store guidelines<sup>1</sup> list several quality criteria. Testing can be done through alpha and beta publications, visible to a private team of testers. However, the store does not check the quality criteria during publication phase. Applications may be published in minutes, even if they do not provide minimum quality requirements. The quality guidelines include recommendations for user interaction, functionality, compatibility, performance, security, publishing, and testing.

The iOS publishing procedure is more reliable. It provides an intermediary environment called TestFlight, which might be used for testing and homologation. There is an automatic quality check before deploying to TestFlight. Publishing to the iOS app

---

<sup>1</sup> <https://developer.android.com/docs/quality-guidelines/core-app-quality>, visited in July 2019

store goes through a manual check by an Apple employee, who checks the correct implementation of quality guidelines<sup>2</sup> for Safety, Performance, Business, Design and Legal Requirements.

An initiative called AQuA<sup>3</sup> proposed a more comprehensive set of guidelines. App Quality Alliance (AQuA) is a group of volunteer developers concerned with the quality of mobile applications. The guidelines are deeply concerned with Resource Utilization (e.g. care with battery life), Fault Tolerance (e.g. handling exceptions when connection is unavailable), Data Persistence (e.g. implement pause, suspend and resume capabilities), Functional Correctness (e.g. grant the correct implementation of calculations), Confidentiality (e.g. do not store sensitive data), among others.

## **1.2 Motivation and Objective**

The growth of mobile application market, in contrast to the seeming lack of comprehensive quality guidelines, makes room for new studies about context-specific quality concerning mobile applications. Several studies propose sets of quality characteristics for apps. (Spriestersbach & Springer, 2004) list common challenges of mobile applications; (Corral et al., 2014) use app store guidelines as a basis to obtain a list of quality characteristics; (Idri, Bachiri, & Fernández-Alemán, 2016) focus on the quality of pregnancy monitoring apps. Meanwhile, none of these studies presents a generic and comprehensive set of quality characteristics for mobile applications.

Context-specific quality characteristics, when known in advance, may guide the development of an application from the specification phase. They may also be used to define evaluation procedures for appraising software already on the market.

The previously presented guidelines are intended to help both development and publication of mobile applications, yet they lack some of their particularities.

This dissertation presents both a main objective and a secondary objective. The main objective is to identify a set of quality characteristics which should be considered when appraising mobile applications. This set shall be used to adapt the ISO/IEC 25010 (ISO/IEC, 2011) quality models for the context of mobile applications. The secondary objective is to apply the adapted quality models in the development of an appraisal

---

<sup>2</sup> <https://developer.apple.com/app-store/review/guidelines/>, visited in July 2019

<sup>3</sup> <https://www.appqualityalliance.org/>, visited in July 2019

procedure. This dissertation presents the following research question: **Which quality characteristics are essential for evaluating mobile applications?**

### 1.3 Methodology

Figure 1.1 presents this dissertation's methodology. Each step will be briefly described below.

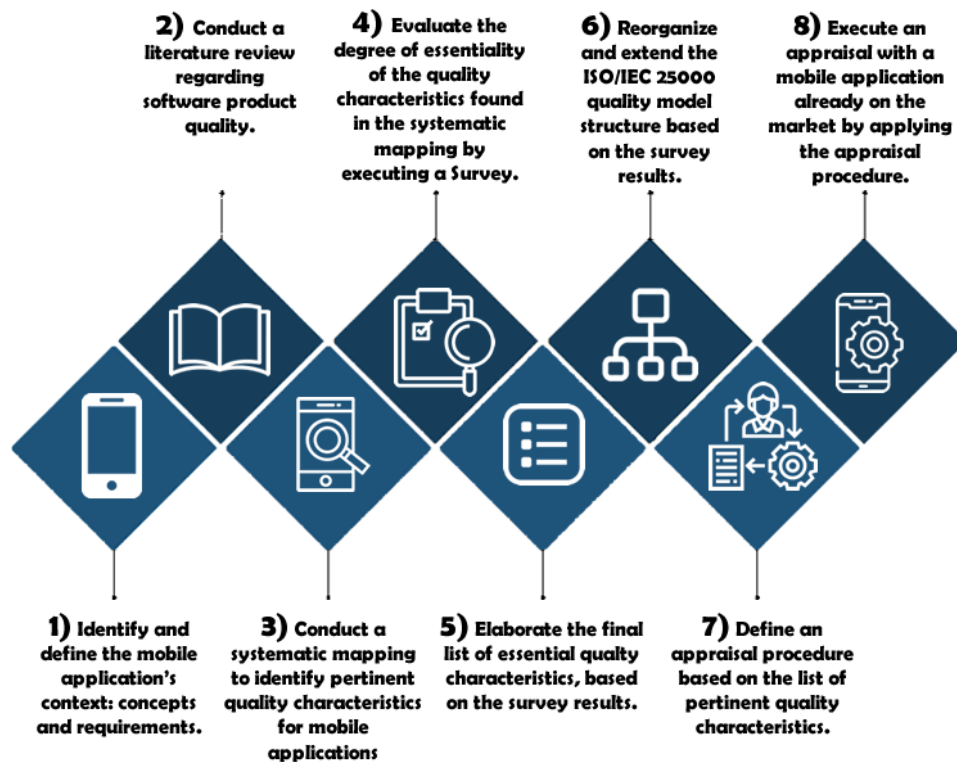


Figure 1.1 – Methodology

- 1) The context of mobile applications was investigated regarding concepts and requirements.
- 2) The literature concerning software quality was also investigated, comprising the evolution of quality models and standards, the identification of evaluation methods and a review of particularization studies for specific products.
- 3) A systematic mapping was conducted in order to identify quality characteristics for the context of mobile applications present in technical literature.
- 4) The results of the systematic mapping were confirmed by executing a survey with mobile application users, regarding the identified quality characteristics.
- 5) Based on the survey results, a final list of essential quality characteristics for the context of mobile applications was elaborated.

- 6) The ISO/IEC 25010 quality models (ISO/IEC, 2011) were adapted to contain only essential quality characteristics for mobile applications.
- 7) An appraisal procedure, based on the list of essential quality characteristics and on the QPS (Rocha et al., 2017) method was defined. QPS is a Brazilian model for appraising software products.
- 8) A mobile application available in the market was appraised using the previously developed procedure.

## **1.4 Dissertation's Organization**

**Chapter 1** introduces the subject of this study, motivation, research question, methodology and dissertation's organization.

**Chapter 2** presents the main concepts of software quality, including the evolution of models and standards, with an emphasis on ISO/IEC 25000 (ISO/IEC, 2011) and on QPS reference model (Rocha et al., 2017) for appraising software products. It also presents examples of studies which particularizes quality models for specific domains.

**Chapter 3** presents a systematic mapping concerning the quality of mobile applications. The literature is reviewed in order to identify a preliminary list of most cited quality characteristics in previous studies about the quality of mobile applications.

**Chapter 4** presents the execution and results of a survey, which confirms the results of the systematic mapping by questioning mobile users about their opinion on quality characteristics in the context of mobile applications.

**Chapter 5** presents an appraisal procedure based on QPS method, on the systematic mapping and survey results. It also presents the application of the appraisal procedure in order to appraise a mobile application.

**Chapter 6** presents the dissertation's conclusion, listing the obtained results, bibliographic production and future work.

## 2 Software Product Quality

This chapter presents the theoretical foundations for the development of this dissertation. It presents the quality models history, emphasizing ISO/IEC 25000 (ISO/IEC, 2011) and the QPS reference model (Rocha et al., 2017) due to their importance in the context of this work. This chapter also presents several studies on quality models particularizations for specific domains.

### 2.1 Software Quality

Product quality is closely related to the conformity with requirements as well as the satisfaction when using a product which is related to its performance and the absence of faults and failures (Guerra & Colombo, 2009). Software must satisfy the user's needs and behave accordingly. This is a subjective judgment, difficult to make without a proper quality appraisal procedure.

Quality models present a set of quality characteristics for software products. However, not every quality characteristic is essential to every type of software. Different uses, requirements, hardware and other inherent characteristics of software products makes some quality characteristics more important in specific contexts. The essential quality characteristics of a specific type of software, when known in advance, can guide the development from the specification phase and ensure the development of products with higher quality.

Due to the advantage of knowing an essential quality characteristic in advance, several studies in literature particularize quality models in order to better understanding of their specific contexts. Some of these studies will be presented in section 2.3.

### 2.2 Historical Quality Models for Software Product

Quality models, intended for software products, propose a hierarchy of quality characteristics and define criteria to be considered during the development or software appraisals, from the stakeholder's point of view. The concern with software quality began in the 70s and continues a relevant subject nowadays. From a historical point of view, two quality evaluation models are particularly important: McCall (McCall, 1977) and Boehm (Boehm, 1978) models.



The McCall model (McCall, 1977) presented a four-level hierarchy. The first level listed three important points of view when evaluating the product quality: Operation, Revision and Transition. Factors, the second level, are associated with these points of view, which are determining characteristics for the software quality. A third level, Criteria, is associated with Factors and can be judged, defined and measured. The fourth and last level defines measures, which permit the measurement of a certain Criteria in a software product. Factors reflect the user's needs and Criteria reflect the developers' perspective. The model contains 11 factors and 23 criteria. Figure 2.1 shows their relation.

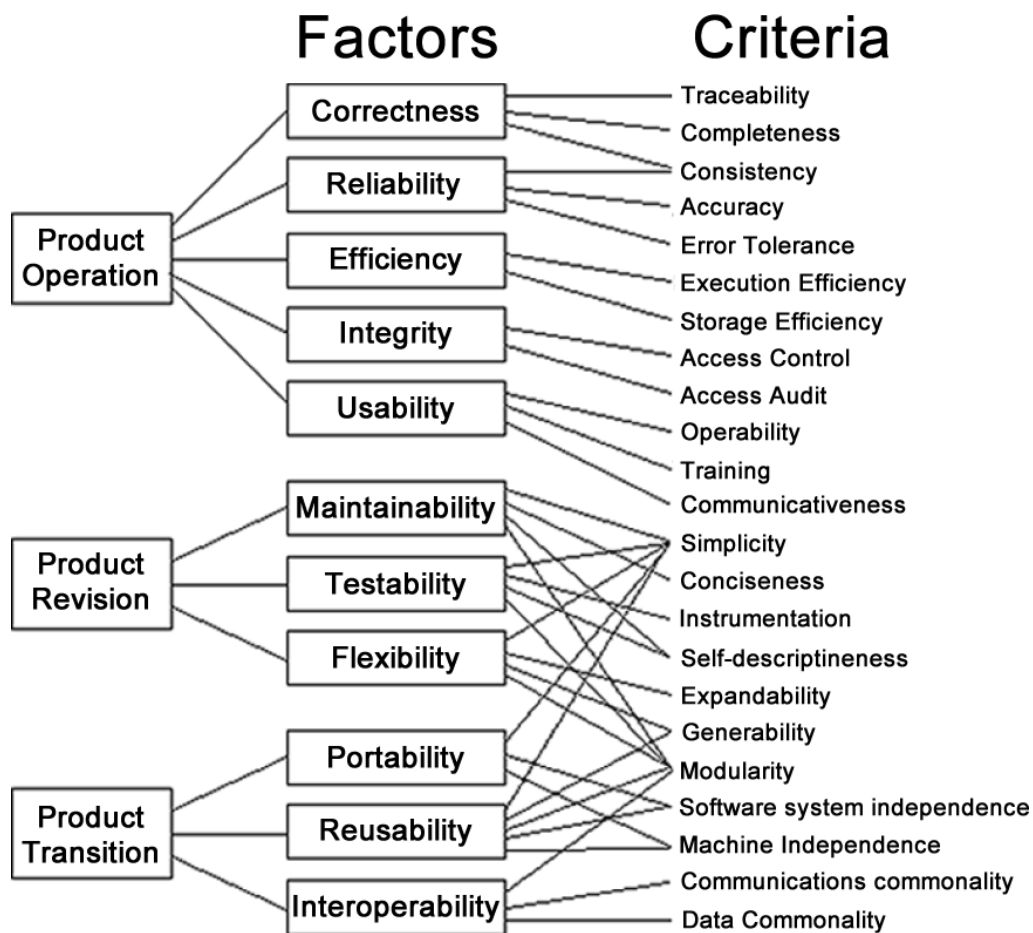


Figure 2.1 – McCall Quality Model (McCall, 1977)

The Boehm model (Boehm, 1978) aimed to evaluate software quality based on quality characteristics and source code metrics. The characteristics are divided in a three-level hierarchy, in which the first level defines three main product uses: As-is Utility, Maintainability and Portability. Figure 2.2 summarizes the model structure.

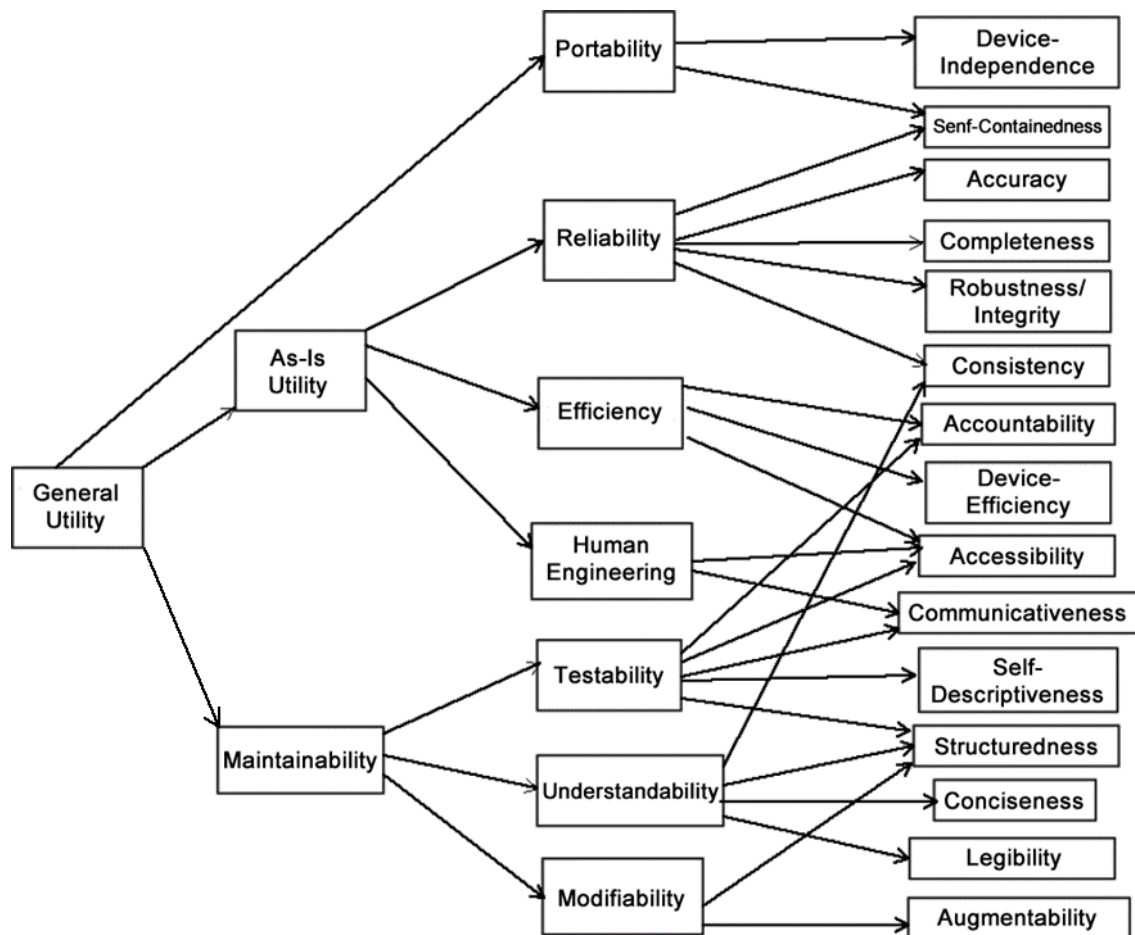


Figure 2.2 – Boehm Quality Model (Boehm, 1978)

## 2.3 International Standards for Product Quality

ISO/IEC 9126 (ISO/IEC, 2001) is the first software quality international standard, predecessor of the current ISO/IEC 25000 series. It presented internal, external and quality in use metrics, besides two quality models. The external and internal Quality Model defined a hierarchy of characteristics and sub-characteristics. Figure 2.3 shows the external and internal Quality Model. The Quality in Use Model did not present sub-characteristics, only the quality characteristics Effectiveness, Productivity, Safety and Satisfaction.

ISO/IEC 9126 was revised and updated, giving place to the ISO/IEC 25000 standards, also known as SQuaRE (System and Software Quality Requirements and Evaluation), which contains a set of standards related to software quality. The numbering of each standard defines its subject: 2500m defines management, 2501n defines quality models, 2502n defines measurements, 2503n defines requirements, 2504n defines quality evaluation methods and both 2505n and 2506n are extensions. The standards whose

subjects are relevant to the context of this work will be presented in the following subsections.

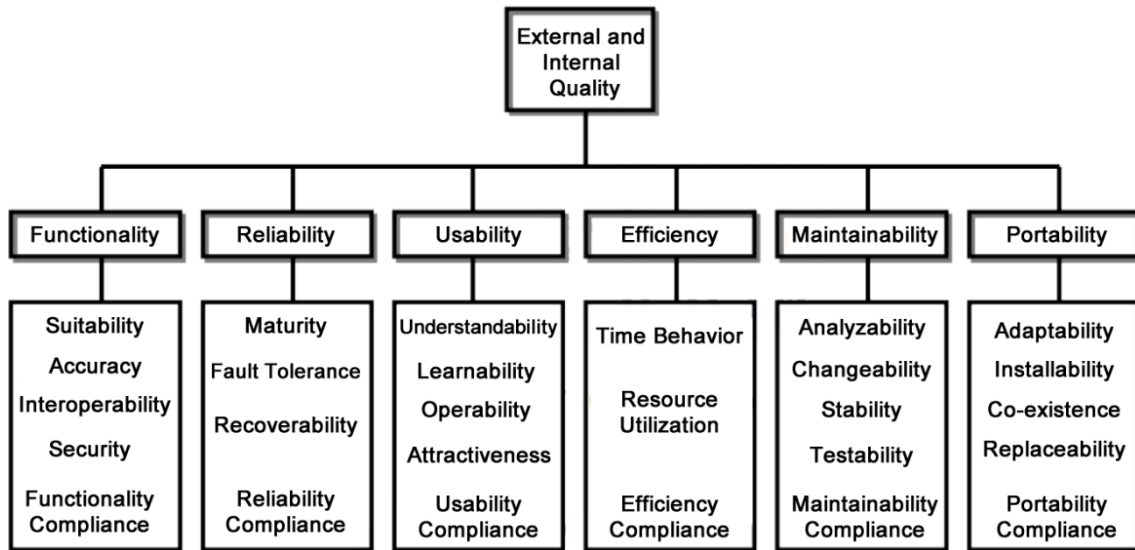


Figure 2.3 – ISO/IEC 9126 External and Internal Quality Model (ISO/IEC, 2001)

### 2.3.1 International Standard ISO/IEC 25010

In 2011, ISO/IEC 9126 was revised and incorporated into the new ISO/IEC 25010. Some characteristics were moved, others were renamed, some were added and certain sub-characteristics rose in the hierarchy and became characteristics.

Software Quality is defined as “the degree to which the system satisfies the stated and implied needs of its various stakeholders and thus provides value” (ISO/IEC, 2011) . The standard defines three types of stakeholders: primary (interact with the product), secondary (provide support, service and management) and indirect (influenced by the results, even without interacting with the product). The standard also defines two quality models:

- Product Quality Model - integrates software static and dynamic aspects. Figure 2.4 presents the model and Table 2.1 presents the definitions of each characteristic and sub-characteristic from this model;
- Quality in Use Model - considers the result of the interactions when a software product is used in a given context. Figure 2.5 presents the model and Table 2.2 presents the definitions of each characteristic and sub-characteristic from this model.

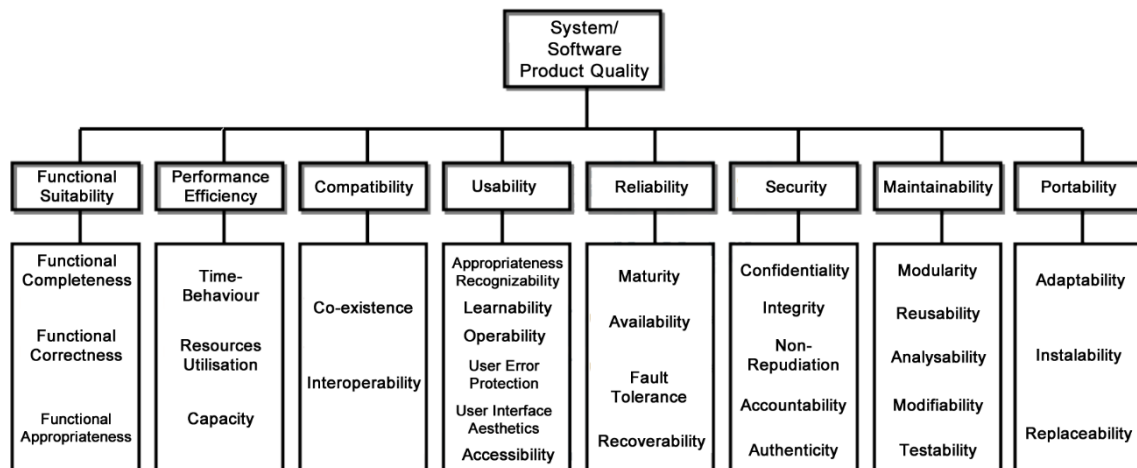


Figure 2.4 – ISO/IEC 25010 Product Quality Model (ISO/IEC, 2011)

Table 2.1 – Characteristics and Sub-characteristics in the Product Quality Model

Characteristics & Sub-characteristics		Definition
<b>Functional Suitability</b>		Degree to which a product or system provides functions that meet stated and implied needs when used under specified conditions.
<b>Sub-characteristics</b>	<b>Functional Completeness</b>	Degree to which the set of functions covers all the specified tasks and user objectives.
	<b>Functional Correctness</b>	Degree to which a product or system provides the correct results with the needed degree of precision.
	<b>Functional Appropriateness</b>	Degree to which the functions facilitate the accomplishment of specified tasks and objectives.
<b>Performance Efficiency</b>		Performance relative to the amount of resources used under stated conditions.
<b>Sub-characteristics</b>	<b>Time Behaviour</b>	Degree to which the response and processing times and throughput rates of a product or system, when performing its functions, meet requirements.
	<b>Resource Utilization</b>	Degree to which the amounts and types of resources used by a product or system, when performing its functions, meet requirements.
	<b>Capacity</b>	Degree to which the maximum limits of a product or system parameter meet requirements.
<b>Compatibility</b>		Degree to which a product, system or component can exchange information with other products, systems or components, and/or perform its required functions, while sharing the same hardware or software environment.
<b>Sub-characteristics</b>	<b>Co-Existence</b>	Degree to which a product, system or component can exchange information with other products, systems or components, and/or perform its required functions, while sharing the same hardware or software environment.
	<b>Interoperability</b>	Degree to which a product can perform its required functions efficiently while sharing a common environment and resources with other products, without detrimental impact on any other product.

Characteristics & Sub-characteristics		Definition
	<b>Usability</b>	Degree to which a product or system can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use.
<b>Sub-characteristics</b>	<b>Appropriateness Recognizability</b>	Degree to which users can recognize whether a product or system is appropriate for their needs.
	<b>Learnability</b>	Degree to which a product or system can be used by specified users to achieve specified goals of learning to use the product or system with effectiveness, efficiency, freedom from risk and satisfaction in a specified context of use
	<b>Operability</b>	Degree to which a product or system has attributes that make it easy to operate and control.
	<b>User Error Protection</b>	Degree to which a system protects users against making errors.
	<b>User Interface Aesthetics</b>	Degree to which a user interface enables pleasing and satisfying interaction for the user.
	<b>Accessibility</b>	Degree to which a product or system can be used by people with the widest range of characteristics and capabilities to achieve a specified goal in a specified context of use.
	<b>Reliability</b>	Degree to which a system, product or component performs specified functions under specified conditions for a specified period of time.
<b>Sub-characteristics</b>	<b>Maturity</b>	Degree to which a system, product or component meets needs for reliability under normal operation.
	<b>Availability</b>	Degree to which a system, product or component is operational and accessible when required for use.
	<b>Fault Tolerance</b>	Degree to which a system, product or component operates as intended despite the presence of hardware or software faults.
	<b>Recoverability</b>	Degree to which, in the event of an interruption or a failure, a product or system can recover the data directly affected and re-establish the desired state of the system.
	<b>Security</b>	Degree to which a product or system protects information and data so that persons or other products or systems have the degree of data access appropriate to their types and levels of authorization.
<b>Sub-characteristics</b>	<b>Confidentiality</b>	Degree to which a product or system ensures that data are accessible only to those authorized to have access.
	<b>Integrity</b>	Degree to which a system, product or component prevents unauthorized access to, or modification of computer programs or data.
	<b>Non-repudiation</b>	Degree to which actions or events can be proven to have taken place, so that the events or actions cannot be repudiated later.
	<b>Accountability</b>	Degree to which the actions of an entity can be traced uniquely to the entity.
	<b>Authenticity</b>	Degree to which the identity of a subject or resource can be proved to be the one claimed.
	<b>Maintainability</b>	Degree of effectiveness and efficiency with which a product or system can be modified by the intended Maintainers.
<b>Sub Sub-characteristics</b>	<b>Modularity</b>	Degree to which a system or computer program is composed of discrete components such that a change to one component has minimal impact on other components.
	<b>Reusability</b>	Degree to which an asset can be used in more than one system, or in building other assets.
	<b>Analysability</b>	Degree of effectiveness and efficiency with which it is possible to assess the impact on a product or system of an intended change to one or more of its parts, or to diagnose a product for deficiencies or causes of failures, or to identify parts to be modified.

Characteristics & Sub-characteristics		Definition
	<b>Modifiability</b>	Degree to which a product or system can be effectively and efficiently modified without introducing defects or degrading existing product quality.
	<b>Testability</b>	Degree of effectiveness and efficiency with which test criteria can be established for a system, product or component and tests can be performed to determine whether those criteria have been met.
<b>Portability</b>		Degree of effectiveness and efficiency with which a system, product or component can be transferred from one hardware, software or other operational or usage environment to another.
<b>Sub-characteristics</b>	<b>Adaptability</b>	Degree to which a product or system can effectively and efficiently be adapted for different or evolving hardware, software or other operational or usage environments.
	<b>Instalability</b>	Degree of effectiveness and efficiency with which a product or system can be successfully installed and/or uninstalled in a specified environment.
	<b>Replaceability</b>	Degree to which a product can replace another specified software product for the same purpose in the same environment.

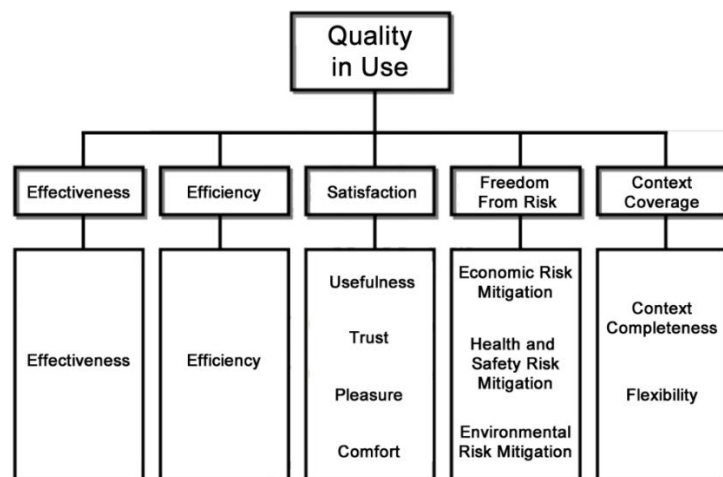


Figure 2.5 – ISO/IEC 25010 Quality in Use Model (ISO/IEC, 2011)

Table 2.2 – Characteristics and Sub-characteristics in the Quality in Use Model

Characteristics & Sub-characteristics		Definition
<b>Effectiveness</b>		Accuracy and completeness with which users achieve specified goals.
<b>Efficiency</b>		Resources expended in relation to the accuracy and completeness with which users achieve goals.
<b>Satisfaction</b>		Degree to which user needs are satisfied when a product or system is used in a specified context of use.
<b>Sub-characteristics</b>	<b>Usefulness</b>	Degree to which a user is satisfied with their perceived achievement of pragmatic goals, including the results of use and the consequences of use.
	<b>Trust</b>	Degree to which a user or other stakeholder has confidence that a product or system will behave as intended.
	<b>Pleasure</b>	Degree to which a user obtains pleasure from fulfilling their personal needs.

Characteristics & Sub-characteristics		Definition
	<b>Comfort</b>	Degree to which the user is satisfied with physical comfort.
	<b>Freedom from risk</b>	Degree to which a product or system mitigates the potential risk to economic status, human life, health, or the environment.
<b>Sub-characteristics</b>	<b>Economic risk mitigation</b>	Degree to which a product or system mitigates the potential risk to financial status, efficient operation, commercial property, reputation or other resources in the intended contexts of use.
	<b>Health and safety risk mitigation</b>	Degree to which a product or system mitigates the potential risk to people in the intended contexts of use.
	<b>Environmental risk mitigation</b>	Degree to which a product or system mitigates the potential risk to property or the environment in the intended contexts of use.
	<b>Context coverage</b>	Degree to which a product or system can be used with effectiveness, efficiency, freedom from risk and satisfaction in both specified contexts of use and in contexts beyond those initially explicitly identified.
<b>Sub-characteristics</b>	<b>Context completeness</b>	Degree to which a product or system can be used with effectiveness, efficiency, freedom from risk and satisfaction in all the specified contexts of use.
	<b>Flexibility</b>	Degree to which a product or system can be used with effectiveness, efficiency, freedom from risk and satisfaction in contexts beyond those initially specified in the requirements.

### 2.3.2 International Standard ISO/IEC 25051

ISO/IEC 25051 (ISO/IEC, 2014) defines a set of quality requirements and conformity evaluation instructions for each quality characteristic in ISO/IEC 25010. The standard also presents a list of generic requirements which shall be relevant to any software, as long as the quality characteristic is pertinent in the given context.

The standard defines requirements for the product description, user documentation and software. It specifically directs the set of requirements to the evaluation of RUSP, Ready to Use Software Product; however, mobile applications are included in this classification. Table 2.3 to Table 2.5 contains examples of requirements.

Table 2.3 – Examples of Product Description Requirements

Characteristic	Product Description Requirements
<b>Usability</b>	The product description shall specify the specific knowledge required for the use and operation of the software.
<b>Security</b>	The product description shall contain, as applicable, statements on Security, taking into account Confidentiality, Integrity, Non-repudiation, Accountability and Authenticity, written such that verifiable evidence of compliance can be demonstrated, based on ISO/IEC 25010.
<b>Portability</b>	The product description shall provide information on the installation procedure.

Table 2.4 – Examples of User Documentation Requirements

Characteristic	User Documentation Requirements
<b>Usability/Operability</b>	If user documentation is not provided in printed form, the documentation shall indicate whether it can be printed, and if so, how to obtain a printed copy.
<b>Security</b>	The user documentation shall provide the information necessary to identify the level of security managed by the software for each data managed by the user.
<b>Compatibility</b>	The user documentation shall provide the necessary information to identify the compatibility to use the software.

Table 2.5 – Examples of Software Requirements

Characteristic	Software Requirements
<b>Usability</b>	The user shall recognize whether the product or system is appropriate for its needs based on the product description or after first manipulation.
<b>Reliability</b>	The software shall perform in accordance with the Reliability features defined in the user documentation.
<b>Security</b>	The software shall have the ability to manage access right management regarding security features.

## 2.4 Software Quality Evaluation and Certification

This section presents national and international initiatives for evaluating the quality of software products based on international standards.

### 2.4.1 National Context

Two national initiatives, although deprecated, were invaluable for the Brazilian software industry. MEDE-PROS (Software Product Quality Assessment Method) guide (Guerra & Colombo, 2009) developed by CTI/CenPRA, proposes checklists to be used in software product appraisals, based on the quality characteristics in ISO/IEC 9126. The checklists refer to elements such as installation, user documentation, product description, packaging and uninstalling. Depending on the product specification, the checklist items might be freely included or excluded. Table 2.6 presents some examples of item checklists.

CERTICS (Alves et al., 2014) was developed in order to be a Brazilian certification for evaluating whether a software product was the result of technological innovation. It was originally intended to facilitate the preference for national products in public procurements and direct the selection of software products. CERTICS evaluated both the software product and the used processes but did not evaluate the organization itself. The CERTICS methodology defined a variety of expected results for the software



product, with which it was possible to characterize if some technological innovation was involved.

Table 2.6 – Examples of checklist items in MEDE-PROS guide

Checklist Title	Item
<b>Completeness of documentation for installation</b>	The documentation provides installation instructions to be read and understood before the installation procedure.
	The documentation is clear about the procedures which should be performed.
<b>User Documentation Completeness: Product Identification</b>	The name of the software is available.
	The version and creation date are available.
<b>Packaging: Usability and Intelligibility</b>	Contains easy to read texts, with decent size of letters.
	Contains colors which make it easy to read and understand information.
<b>Uninstalling</b>	The product has an uninstall procedure, which may be automated or manual.
	The product's messages display the progress of the task.

#### 2.4.2 International Context

The Spanish Association for Standardization and Certification (AENOR) identified several studies concerning process evaluation, a few studies on the evaluation of software product and no studies related to evaluation and certification of software products. Consequently, AENOR decided to elaborate a new certification based on ISO/IEC 25000 standards (Rodríguez et al., 2015). AENOR defined an ecosystem for appraising and certifying software product quality. The ecosystem consists of:

- Companies interested in a product quality evaluation;
- AENOR's certifying organ, which comprises more than 20 years of experience in product quality audits;
- AQC Lab, the first certified laboratory for software appraisal based on ISO/IEC 25000 standards. The laboratory implemented an appraisal process based on ISO/IEC 25040, tools to automatize measurement and a quality model with characteristics and measures, complementary to ISO/IEC 25010;
- Expert consultants in software quality;

- Companies involved in the development of software measurement tools. Among these tools, we highlight Kiuwan and SonarQube.

The certification starts with a pre-assessment, based on ISO/IEC 25010. The interested companies should submit a product appraisal request to AQC Lab, which is responsible for the pre-assessment. The company itself must indicate which quality characteristics would be appraised. Give the results of the pre-assessment, the organization might opt for either improving the product or using them, in case it suits the company's expectations. In the latter case, the company might request an AENOR certification.

### 2.4.3 QPS

Recently, a newer initiative for conducting software products evaluation arose, the QPS (Software Product Quality) reference model. QPS (Rocha et al., 2017) is a Brazilian model for appraising software products. The QPS structure comprises four dimensions:

- **Organizational Dimension:** A company in charge of a product in the market shall provide user documentation, organizational processes and the processes' attributes related to the implementation of Service Dimension and Software Engineering Dimension.
- **Software Engineering Dimension:** A company in charge of a product in the market shall provide corrective, evolutive and adaptive maintenance processes (if pertinent).
- **Service Dimension:** A company in charge of a product in the market shall provide a customer support service.
- **Product Quality Dimension:** A software product in the market shall contain description, user documentation and quality characteristics, which may be: (I) essential for any product; and (II) characteristics which shall be present in certain products to meet their specific features.

The QPS model is based on the principles of the continuous assessment, by presenting results in three levels of recognition: gold, silver and bronze. Dimensions are based on a series of international standards, considering the extent of the appraisal and the necessity of maintaining it in conformity with the existing quality perspective:

- **Organizational Dimension** - ISO/IEC/IEEE 12207 and ISO/IEC 33000;

- **Software Engineering Dimension** - ISO/IEC 12207 and ISO/IEC 33000;
- **Service Dimension** - ISO/IEC 20000 and ISO/IEC 33000;
- **Product Quality Dimension** - ISO/IEC 25000.

The model distinguishes either products in launch phase or existing products already on the market. For existing products on the market, the model evaluated both the existence and the execution of processes. The evaluation assigns one of three levels: gold, silver or bronze, according to the degree of coverage of the expected results. The achievement of levels is incremental: in order to achieve a level, it is mandatory to meet every requirement stated at the previous levels. Figure 2.6 presents an overview of the QPS model structure.

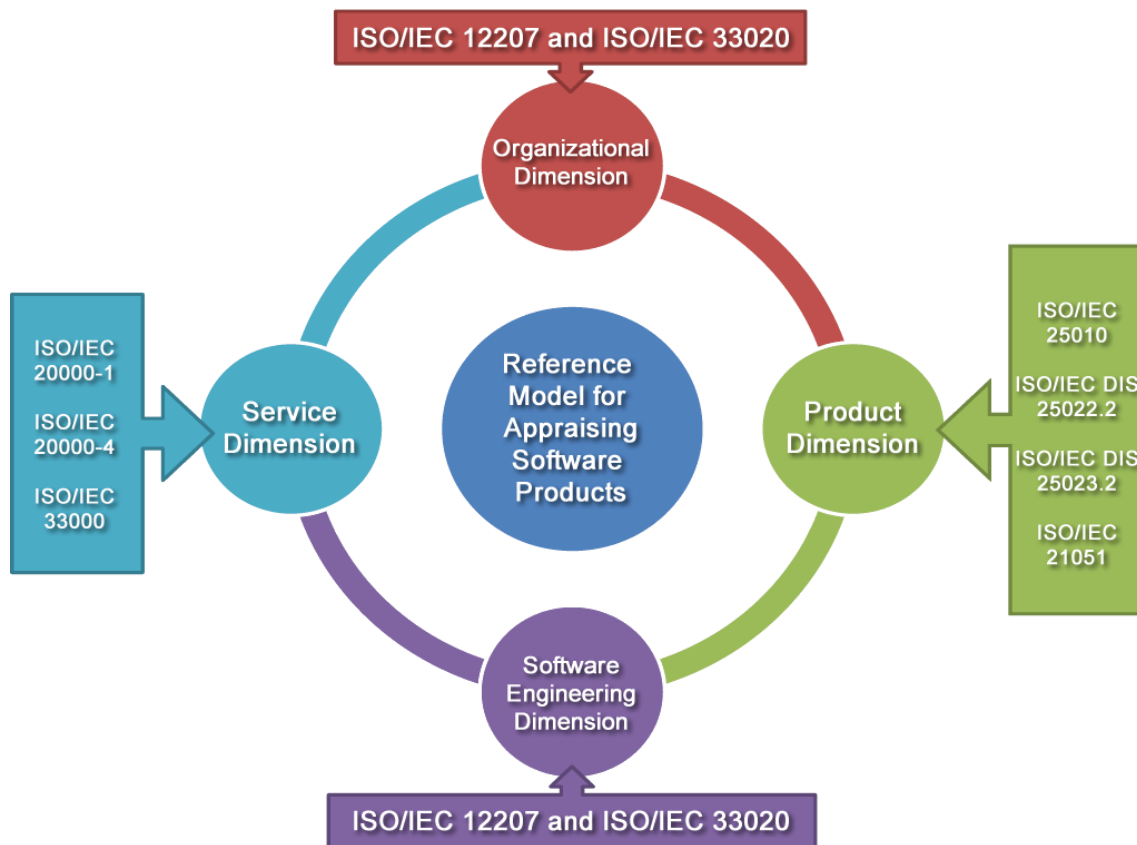


Figure 2.6 – Overview of QPS Model and the relation with international standards

The Product Quality Dimension evaluates if the product provides an available description for possible users and buyers, as well as user documentation. The product should also conform to a subset of measures related to quality characteristics: operational consistency of messages, existence of undo option, user aesthetics, access control,

availability of analysis documentation, documentation support for testing and availability of test cases.

The following results shall be evidenced at silver level:

- Conducting a survey with companies which acquired and uses the product, in a period of at most 12 months before the evaluation date;
- The results of the survey;
- Evidence that the results were analysed and that the necessary measures have been taken in order to improve the product where relevant.

The following results shall be evidenced at gold level:

- Identification of the products' context-specific quality requirements;
- Existence of procedure for evaluating the level of understanding of the product regarding the quality requirements.

The context-specific requirements should be identified considering the product's specific characteristics as well as the needs of every user type (e.g. final users, indirect users, maintainers, content creators, etc.)

The product quality dimension is directly related to the results of this work. The quality characteristics considered in a QPS appraisal may be general-purpose, but the model also intends to evaluate specific quality characteristics in the gold level.

## **2.5 Particularizations of Quality Models**

Quality models are general-purpose and made to be applied to any kind of software product. However, not every quality characteristic is appropriate in all cases and some may be more relevant depending on context, hardware and stakeholders. This generality makes it necessary particularize quality models for different scenarios. We shall present some examples of studies which particularize quality models.

There is a series of studies (Franch & Carvallo, 2003) suggesting a methodology, tools and criteria for selecting COTS (Commercial off-the-shelf) for organizations. These studies present a methodology consisting of the decomposition of ISO/IEC 9126 quality characteristics, until each attribute became measurable. The methodology contains seven steps, ranging from the analysis of quality characteristics to the definition of metrics.

In (Barney & Wohlin, 2009) a case study is reported in which a questionnaire was elaborated with the quality characteristics in ISO/IEC 9126 and some other models found in literature. Several employees from Ericsson with different functions and involved in different projects were requested to distribute 1000 points among the characteristics they considered the most important. The objective of the study was to identify different stakeholder's points of view.

A survey was carried out in (Haigh, 2010), with the purpose of identifying opinions about the importance of quality characteristics in the Boehm quality model. The respondents were students and ex-students from an unnamed MBA program in the USA, each performing different roles in software projects. Users prioritized integrity, portability and usability. Developers prioritized maintainability and testability. Managers prioritized accuracy.

In (Trienekens et al., 2010) a case study involving the management system of warships is described. In this context, two stakeholders were spotted: military specialist operators and the quality assurance team. Problems such as unpredictable execution time, restricted reliability and subjectivity of the quality specification led the team to seek support of a quality model, in order to evaluate the key characteristics for the system to operate satisfactorily. The stakeholders discussed about the definitions in ISO/IEC 9126 and reduced it to a shorter list of quality characteristics, specific for their functionalities.

The study of (Fahmy et al., 2012) had the purpose of determining a quality model for software of "e-Books", a mini laptop with academic applications. It listed 35 criteria considered important for this type of software and related each of them to five quality characteristics in ISO/IE 9126: Functionality, Reliability, Usability, Efficiency and Portability. The most related characteristic was Usability.

In (M. Kim et al., 2017) is described a study with the purpose of determining a quality model for IoT (Internet of Things) based on ISO/IEC 9126 and ISO/IEC 25010. The security approach in ISO/IEC 9126 was considered inappropriate, which motivated the use of both standards. The study lists general characteristics of IoT applications: mobility, connectivity, resource limitations and need for security. As a conclusion, metrics were developed for five quality characteristics: Functionality, Reliability, Efficiency, Portability (from ISO/IEC 9126) and Security (from ISO/IEC 25010).

The studies briefly described above evidence the importance and necessity of identifying context-specific quality characteristics.

## **2.6 Conclusion**

This chapter presented two historical quality models McCall and Boehm, as well as international standards ISO/IEC 9126 and its successor ISO/IEC 25010, which provides two quality models: product quality and quality in use.

Regarding evaluation and certification, the deprecated national initiatives MEDE-PROS guide and CERTICS were briefly introduced, also AENOR on the international scene. QPS, a newer initiative for appraisal of software products was presented. QPS is strongly related to the results of this dissertation.

We presented several studies which particularize quality models for different domains, in order to identify which characteristics are essential for specific types of software.

The next chapter will present a systematic mapping which aims to characterize the current state of researches into the quality of mobile applications.

# 3 Systematic Mapping

This chapter presents the protocol and the results of the systematic mapping concerning the quality of mobile applications. Section 3.1 presents the review process. Section 3.2 presents the research goal. Section 3.3 describes the planning phase. Section 3.4 describes the execution phase. Section 3.5 describes the analysis phase. Section 3.6 discusses the results. Section 3.7 lists the threats to validity. Section 3.8 presents the conclusion.

## 3.1 Research Method

This systematic mapping aims to characterize the current state of investigations into the quality of mobile applications. A systematic review is a means of identifying, evaluating and interpreting the available research data related to a research question, topic area, or phenomenon. The main purpose of conducting a Systematic Review is to gather evidence on which to base conclusions (Kitchenham & Charters, 2007). A Systematic Mapping (SM) adopts the same rigor and maps the available evidence when no conclusions can be reached (Petersen et al., 2015). Biolchini et al. (2005) propose a process for conducting Systematic Reviews (Biolchini et al., 2005), which consists of four phases, as showed in Figure 3.1:

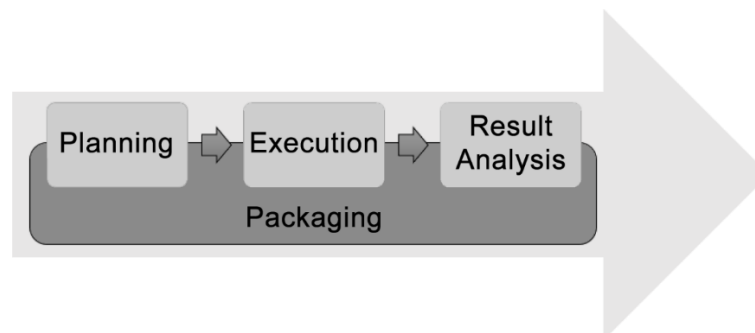


Figure 3.1 – Summary of the Systematic Review Process (Biolchini et al., 2005).

- **Planning:** a phase in which the objectives, research questions and methods for the execution and analysis phases are defined. The search string is assembled and the most suitable search engines are selected. A set of inclusion and exclusion criteria is also defined to filter the resulting papers.
- **Execution:** a phase in which the search string is iteratively executed and refined until the set of resulting papers proves to be satisfactory. The previously defined set of inclusion and exclusion criteria are applied to the resulting papers.

- **Result Analysis:** The papers are read and their relevant data are extracted to an extraction form, according to the research questions. Then, all the extracted data are assembled and interpreted to answer these questions.
- **Packaging:** This step is executed throughout the whole process. The objective is to keep all the decisions and collected information documented.

Kitchenham and Charters (2007) also describe the **Reporting** phase, in which the produced results and documentation are made available to potential stakeholders.

### 3.2 Research Goal

Mobile devices and applications became especially popular after the creation of the app stores' distribution model (Cortimiglia et al., 2011). Since then, the quantity of available mobile applications has increased significantly every year. The variety of apps and the ease of their installation through the app stores make it essential to develop apps of distinctive quality, since it poses no difficulty for users finding and installing competing mobile applications with the same functionalities. Although the app stores contains publication guidelines, they are not broad enough to prevent the submission of mobile applications not in conformity with the users' expectations. Almost every app store permits the publication of unfinished mobile applications, except for the iOS app store, which may reject a submission if it did not follow some basic guidelines.

The main objective of this SM is to identify which quality characteristics are most pertinent in the context of mobile applications. The quality of a system is the degree to which the system satisfies the stated and implied needs of its various stakeholders, and thus provides value (ISO/IEC, 2011). Table 3.1 presents the aim of this study, using the GQM paradigm (Basili et al., 1994).

Notice that the systematic mapping intends to identify *pertinent* quality characteristics, even though the objective of the dissertation, as described in the introduction, is to propose a set of *essential* quality characteristics. The adjectives “pertinent” and “essential” are not treated as synonyms. The pertinent quality characteristics are those somehow related to the context of mobile applications, but not necessarily essential. Afterwards, a survey will be conducted in order to identify which pertinent quality characteristics are also essential.



Table 3.1 – Research goal, according to GQM paradigm

Research Goal	
Analyze	Quality characteristics
For the purpose of	Characterizing
With respect to	Pertinence
From the point of view of	Software engineering researchers
In the context of	Mobile applications

### 3.3 Planning

This section presents the protocol for the planning phase. This phase defines the research questions, the search strategy and the inclusion and exclusion criteria.

#### 3.3.1 Research Questions

ISO/IEC 25010 (ISO/IEC, 2011) contains two quality models, each with a hierarchy of characteristics and sub-characteristics. At first, we did not intend to bind the analysis of the results to these models, because we were not sure if the papers in the literature would follow the definitions from these standards. Furthermore, we expected to find attributes beyond those present in these quality models. Surprisingly, the preliminary execution of the search string returned a diversity of papers about the quality of mobile applications that indeed used ISO/IEC 9126 or ISO/IEC 25010 (ISO/IEC, 2011), even though the search string did not explicitly include them.

The original objective of the mapping was to list pertinent quality characteristics of mobile applications independently of existing quality models, but it seemed clear that the result would contain many occurrences of quality characteristics from the two models in ISO/IEC 25010. Due to this observation, two research questions were developed. They are present in Table 3.2.

Table 3.2 – Research Questions

	Research Question	Rationale
RQ1	Which quality characteristics from ISO/IEC 25010 are identified as pertinent in the context of mobile applications?	The answer indicates the ISO/IEC 25010 characteristics considered pertinent in several researches about quality of mobile applications present in literature.
RQ2	Which quality characteristics are not present in ISO/IEC 25010, but are identified as pertinent in the context of mobile applications?	The answer indicates the characteristics considered pertinent in several researches about quality of mobile applications present in the literature, but not covered by ISO/IEC 25010.

### 3.3.2 Search String

The selection of papers was carried out through a search string, applied in several search engines. The string followed the PICO process (Pai et al., 2004). The subject of the research should be decomposed in four parts: population, intervention, comparison and outcome. There is no comparison in the scope of this work.

Our population are mobile applications. Our intervention are software quality models, quality in use and quality requirements. Our outcome are characteristics, metrics, measures, evaluation criteria and attributes. The “quality requirements” intervention was not considered in the first version of the string, although they could be helpful to infer implied qualities of a software product. So, the string was calibrated once to include it.

- Preliminary search string: (*("mobile application") AND ("software quality" OR "quality model" OR "quality in use") AND ("characteristic" OR "metric" OR "measure" OR "evaluation criteria" OR "attribute")*)
- Final search string: (*("mobile app\*") AND ("software quality" OR "quality model" OR "quality in use" OR "quality requirement") AND ("characteristic" OR "metric" OR "measure" OR "evaluation criteria" OR "attribute")*)

### 3.3.3 Selection of Search Engines

The final search string was executed in five search engines: Scopus<sup>4</sup>, IEEE<sup>5</sup>, Web of Science<sup>6</sup>, Engineering Village<sup>7</sup> and ACM<sup>8</sup>. The first execution occurred in May 2018 and the last in September 2018. These engines were selected due to their high research coverage in software engineering or HCI. Except for Scopus, all the other engines were personally recommended by other researchers.

### 3.3.4 Inclusion and Exclusion Criteria

The execution of the search string in the search engines returned several papers, but not all of them might be suitable for the purposes of the research project. Therefore, inclusion and exclusion criteria were developed in order to refine the initial list of papers. They are listed in Table 3.3 and

---

<sup>4</sup> <http://www.scopus.com>

<sup>5</sup> <http://ieeexplore.ieee.org>

<sup>6</sup> <http://apps.webofknowledge.com>

<sup>7</sup> <http://www.engineeringvillage.com>

<sup>8</sup> <http://dl.acm.org>

Table 3.4.

Table 3.3 – Exclusion Criteria

Exclusion Criteria	
E1	The paper was not published in conferences or journals.
E2	The paper is not available in the internet.
E3	The paper is not in English.
E4	The paper is not about quality in mobile applications.
E5	The paper is nearly equal to another paper by the same authors.
E6	The paper is about design patterns, source code or quality of services.

Table 3.4 – Inclusion Criteria

Inclusion Criteria	
I1	The paper was published in conferences or journals.
I2	The paper is available in the internet.
I3	The paper is in English.
I4	The paper is about quality in mobile applications.

### 3.4 Execution

The search string was executed in the five search engines. The exclusion criteria E1, E2 and E3 were immediately applied to the initial set of papers. The application of the other exclusion criteria depended on further interpretation, so both the author and another researcher read the abstracts of the remaining papers, then decided together on more exclusions, based on criteria E4, E5 and E6. Some papers could only be safely excluded after the full text was read. Table 3.5 shows the number of results before and after the application of exclusion criteria.

Table 3.5 – Search Results

Search Engine	Initial Set	After E1, E2 & E3	After E4, E5 & E6
Scopus	57	45	29
IEEE	27	26	13
Engineering Village	27	19	13
Web of Science	30	23	13
ACM	7	7	3

Most of the papers were found in more than one search engine. In total, 35 papers were selected using the search string and 18 using snowballing (i.e., manually selected from the references of included papers). Table 3.6 lists the exclusions, and Table 3.7 and Table 3.8 list the papers included via search string and snowballing, respectively. Table 3.9 briefly summarizes every included paper. The ID in Table 3.7 and Table 3.8 will be used from now on to reference each paper. Even though Table 3.8 contains the manually

selected papers, they were also individually searched in each engine, as an additional information.

Table 3.6 – Excluded papers and the engines where they were found

	Papers	Exclusion Criteria	Scop.	IEEE	Eng.V	WoS	ACM
1	(Aranha & Borba, 2007)	E4					
2	(Canfora et al., 2016)	E6					
3	(Bezerra et al., 2014)	E4					
4	(de Souza & de Aquino, 2014)	E5					
5	(Potena, 2013)	E4					
6	(Hilwa & Samidi, 2014)	E4					
7	(Cortellessa et al., 2010)	E4					
8	(Hyun & Soo, 2013)	E6					
9	(Ribeiro & Dias-Neto, 2017)	E4					
10	(Rao et al., 2012)	E4					
11	(Rohil & Gupta, 2012)	E4					
12	(Gronli & Ghinea, 2016)	E6					
13	(Hecht et al., 2015)	E6					
14	(Corral & Fronza, 2015)	E4					
15	(Liao et al., 2017)	E4					
16	(Orru et al., 2015)	E6					
17	(Bachiri et al., 2015)	E5					
18	(Hecht et al., 2016)	E6					
19	(Linares-Vásquez et al., 2014)	E6					
20	(Seshasayee et al., 2007)	E4					
21	(Zernadji et al., 2016)	E4					
22	(Ricciardi et al., 2015)	E4					
23	(Syer et al., 2015)	E6					
24	(Naab et al., 2015)	E4					
25	(Catolino, 2018)	E6					

Table 3.7 – Papers included via search string and where they were found

ID	Papers	Scop.	IEEE	Eng.V	WoS	ACM
M1 <sup>9</sup>	(Fauzia et al., 2014)					
M2 <sup>6</sup>	(Idri et al., 2017)					
M3	(Mohsin et al., 2017)					
M4	(Barnett et al., 2015)					
M5	(Idri, Bachiri, & Fernández-Alemán, 2016)					
M6	(Franke et al., 2012)					
M7	(Baloh et al., 2015)					
M8	(Yildiz et al., 2014)					
M9	(Alaa et al., 2013)					
M10	(Kabir et al., 2016)					
M11	(Idri, Bachiri, Fernandez-Aleman, et al., 2016)					
M12	(de Souza & de Aquino, 2015)					
M13	(Fang et al., 2017)					
M14	(Pretel & Lago, 2012)					
M15	(Franke & Weise, 2011)					

<sup>9</sup> Control group papers

ID	Papers	Scop.	IEEE	Eng.V	WoS	ACM
M16	(Hess et al., 2012)					
M17	(Soad et al., 2016)					
M18	(Marinho & Resende, 2012)					
M19	(Lew & Olsina, 2013)					
M20	(Zhenyu Liu et al., 2014)					
M21	(Moumane & Idri, 2017)					
M22	(Liu et al., 2014)					
M23	(Ben Ayed et al., 2016)					
M24	(Moumane et al., 2016)					
M25	(Corral, 2012)					
M26	(Holl & Vieira, 2015)					
M27	(Nayebi et al., 2012)					
M28	(Ryan & Rossi, 2005)					
M29	(Corral et al., 2014)					
M30	(Khalid et al., 2015)					
M31	(Olsina & Lew, 2017)					
M32	(Peischl et al., 2015)					
M33	(Khalid et al., 2016)					
M34	(Grano et al., 2017)					
M35	(Abusair, 2017)					

Table 3.8 – Papers included via Snowballing, a manual search in the search engines

ID	Papers	Scopus	IEEE	Eng.V.	WoS	ACM
M36	(Cortimiglia et al., 2011)					
M37	(Dantas et al., 2009)					
M38	(Spriestersbach & Springer, 2004)					
M39	(de Sá & Carriço, 2008)					
M40	(Huang, 2009)					
M41	(Harrison et al., 2013)					
M42	(Lai, 2015)					
M43	(Lim et al., 2015)					
M44	(La et al., 2011)					
M45	(Gafni, 2009)	Not found				
M46	(Holl & Elberzhager, 2014)					
M47	(Holzinger et al., 2012)					
M48	(Wasserman, 2010)					
M49	(H.-W. Kim et al., 2011)					
M50	(Balagtas-Fernandez & Hussmann, 2009)					
M51	(Zahra et al., 2013)	Not found				
M52	(Hussain & Kutar, 2009)	Not found				
M53	(Savio & Braiterman, 2007)	Not found				

Table 3.9 – Summary of the included papers

ID	Summary
M1	It considers every ISO/IEC 25010 product quality characteristics as basis for defining a new quality model, with 17 metrics (uses GQM). A case study with three mobile applications validates two metrics.
M2	Authors' older studies identified that four of the ISO/IEC 25010 product quality characteristics are the most relevant in pregnancy monitoring mobile applications' requirements. This sequel study identifies that, in the analyzed mobile applications,

ID	Summary
	Reliability, Functional Suitability and Usability are the most covered, while the least covered is Performance Efficiency.
M3	Does a literature review about the quality characteristic Fault Tolerance and proposed a set of six measures. They are validated with experimental studies in real environments, with final users using Android mobile applications.
M4	It presents an architecture model for data-intensive mobile applications, with six concepts that influence quality. They are validated by two case studies in which mobile applications are used in real environments. The paper concludes with a set of recommendations for developers.
M5	It lists requirements for pregnancy monitoring mobile applications, obtained both from literature and from analysis of real mobile applications. The requirements were related to each of the ISO/IEC 25010 quality characteristics and the results were applied to formulas for calculating the impact of the characteristics in each block of requirements. The results indicates a greater impact of Functional Suitability, Reliability, Performance Efficiency and Usability.
M6	It proposes a quality model for mobile applications. The authors base their research on Boehm and ISO/IEC 9126, but do not indicate how the model was conceived. It is validated by a case study that compares two mobile applications and explores the characteristics Data Persistence, Usability and Efficiency.
M7	It proposes a framework for evaluation of mobile learning applications, based on a catalog of requirements proposed by another study. The model is validated by applying it on 21 mobile applications.
M8	Every ISO/IEC 25010 product quality characteristics and sub-characteristics were selected from systematic review, also three quality characteristics for B2C from a previous study, totaling 23 sub-characteristics. A survey was answered by developers with the purpose of indicating which of these are most relevant in mobile development.
M9	It focuses on quality of services, calculated with code metrics related to component design. The quality in use characteristics Efficiency and Effectiveness are presented as important.
M10	It discusses the social characteristics of mobile applications. It defines a process for developing social systems, since requirements elicitation until implementation. The process is validates in two case studies.
M11	Previous researches from the same authors identified four quality characteristics most influenced by pregnancy monitoring mobile applications. This paper continues the research.
M12	It proposes an estimation method for mobile applications in design phase. It conducts a literature review in search of mobile application characteristics. 29 are identified, posteriorly refined to 13.
M13	It proposes a research model to identify antecedents of mobile application adoption. The model is composed by perceived characteristics of innovation, taken from a framework called PCI, and other characteristics taken from an unidentified ISO. The model is validated by tourists who uses real mobile applications and then answer a survey.
M14	It develops a way for capturing the interactions of users with the system and with the environment, with the intention of reducing the noise generated by the context in mobile application tests.
M15	It presents a set of quality characteristics without indicating how the authors concluded their importance. It conducts a case study about the importance of the quality characteristic Data Persistence.
M16	It describes a method for creating business mobile applications, focusing on the usability and in the user experience. These are considered key quality characteristics.
M17	It discusses about the definition of an evaluation method for mobile learning applications. The method proposes a model with quality characteristics, metrics and evaluation criteria. The quality characteristics are based on ISO/IEC 25010 and a set of quality characteristics taken from a catalog. The validation is done by applying the evaluation criteria to three mobile applications.

ID	Summary
M18	It defines a procedure to relate ISO/IEC 25010 quality characteristics to development good practices recommended by UTI and W3C. The quality characteristics are ranked based on the level of relationship with the good practices.
M19	It proposes the use of a framework previously defined by the same authors. It discusses the importance of some quality characteristics and comments about their effect in four mobile applications.
M20	It proposes a testing framework for mobile applications, which comprises Security and Portability. Aspects of these two quality characteristics are discussed. There is no validation.
M21	It compares two frameworks previously developed by the authors, intended to provide an overview and discuss similarities and differences. It also presents correlations between ISO/IEC 9126 and limitations of mobile environments.
M22	It proposes a quality model for testing applications. Requirements that should be tested are related to every ISO/IEC 9126 quality characteristic. There is no validation.
M23	Based on previous researches of the author, two ISO/IEC 25010 quality characteristics are considered most relevant for mobile applications. It suggests that the quality characteristics Effectiveness and Efficiency can be objectively evaluated. Then defines measures and conducts a case study.
M24	It describes a framework and apply it in an experiment with two mobile applications, focusing on the user interface limitation. The focus of the framework is to improve usability.
M25	It is an initial version of <b>M29</b> . It does not present quality characteristics, only the objectives and methodology.
M26	In previous studies, the authors proposed a mobile application's failure pattern classification. This paper conducts a survey with experts, a case study and a comparison between two projects, with the purpose of evaluating if the classification is complete, usable and effective.
M27	It presents a literature review about usability of mobile applications. It lists definitions and evaluation methodologies.
M28	It defines efficiency metrics for mobile applications, and validate them with an empirical research that correlates every couple of metric.
M29	It extracts quality requirements from quality guidelines of six app stores and relates them to every ISO/IEC 25010 quality characteristic. The study used the methodology QFD, with which was identified the degree of the requirements' importance in the users' point of view and their relation with the users' degree of need. This association permits the evaluation of the importance of each quality characteristic.
M30	It analyses 10.000 mobile applications with a tool that detects code bugs. It conducts a case study that compares the results to the app's Play Store stars evaluation. Three kinds of warnings are identified as recurrent in apps with negative reviews.
M31	The authors developed a quality model focused on Trust in previous studies. This paper presents the first version of a questionnaire intended to validate the model, to be applied to primary users and experts.
M32	It describes a medical mobile application developed under limited resources. It discusses some attributes related to security, efficiency, portability and usability; also presents requirements that were decisive when deciding where to publish the application. The study conducts an empirical evaluation in which participants answered two questionnaires and execute system activities.
M33	It interprets the texts of a massive quantity of user reviews in iOS App Store and identifies a resulting list of 12 common user complaints.
M34	It extracts around 288 thousand reviews from many versions of 395 open source Android applications, and then apply sentiment analysis to them. According to the subject of each review, they are related to a "topic" or "user intention", from a taxonomy developed by other authors.
M35	It aims to develop a methodology to help the creation of context-aware mobile environments. It considers many characteristics that affect the user satisfaction, like the server availability. There is no validation.

ID	Summary
M36	It comments about the importance of app stores, some features of Apple's app store and benefits of this model of distribution. There is no validation.
M37	It proposes a list of requirements for testing mobile applications, aiming to improve productivity and effectiveness of the testing process. Questionnaires are applied to developers and testers, in order to find out the mostly used tests, and to identify if testers consider the mobile environment limitations. Some tests are used in a real application.
M38	It relates challenges in the development of mobile applications to ISO/IEC 9126 quality characteristics. Then it indicates which quality characteristics are most affected by these challenges.
M39	It presents guidelines for developing mobile application interfaces, by considering many kinds of interactions and context factors that affects usability. It proposes a methodology in which interaction data is collected and analyzed in search of patterns and design problems. The methodology is validated with three case studies.
M40	It presents a list of IHC challenges during the implementation of mobile applications and devices. There is no validation.
M41	It proposes a usability model for mobile applications, which binds quality characteristics from two other models and puts together "Cognitive Load". It conducts a literature review both to assemble the new model and to find out the frequency of each quality characteristic in other studies.
M42	It wants to identify factors that affects traveler's adoption of tour guide mobile applications. The factors were obtained from a literature review. Interviews were conducted in the streets with 206 travelers, who answered questions about the influence of these factors.
M43	It promotes a survey in many countries, with 10.208 answers, in search of user behavior differences due to country differences. The questions are about the reason for selecting an app, the frequency that they evaluate an app and the reasons for abandoning an app.
M44	It presents a methodology with factors and guidelines influencing the effectiveness of mobile applications. The guidelines are validated with three experiments, each with different scenarios, interaction frequency and complexity of features.
M45	It defines metrics for all the sub-characteristics of usability in ISO/IEC 9126. Each metric is validated by at least four experiments.
M46	It conducts a literature review in search of common faults during the development of mobile applications. It proposes a failure classification, a list of categories of faults and a relationship between them. There is no validation.
M47	It reports experiences of the development of a mobile project for accessing limited access database information. It details possible publication methods, focusing on screen size and resolution differences.
M48	It conducts a survey with developers in order to get to know better about mobile development practices. It presents an overview about this type of development, with tools and good practices.
M49	It investigates reasons that leads users to buy mobile applications. It conducts interviews, which helped to identify seven variables that influences this decision.
M50	A framework is developed, based on a four-step methodology (preparation, collection, extraction and analysis). It captures usability variables directly from code with logs, calculates the measures and presents to the user. The study conducts a proof of concept.
M51	It proposes a quality model for mobile applications, based on literature review and on ISO/IEC 9126. It does not explain the origin of the specific selection of quality characteristics. The results are not validated.
M52	It conducts a literature review about usability in systems and IHC. It defines a set of guidelines given the review results. Questions and metrics are defined with GQM. There is no validation.
M53	It presents a model for the context of mobile interaction and a set of design heuristics for successful mobile interactions. There is no validation of the model.



### **3.5 Analysis of the Results**

The 53 selected papers were analyzed by means of the two research questions mentioned above. The extracted data, which answers the research questions, were collected with the assistance of a form. This section shows the conception of this form, discusses the results of each question, and discusses additional results, based on the interpretation of the extractions.

#### **3.5.1 Data Extraction Form**

As far as possible, the research questions must be objectively answered by the reading and interpretation of each paper. An extraction form (Annex A) was developed to keep the data of each read paper. These data would subsequently be put together and analyzed to answer the research questions.

To answer RQ1, the form contains tables to store information on quality characteristics and sub-characteristics of both quality models from ISO/IEC 25010, which might be described as relevant in the paper. To answer RQ2, the form contains a table to retain information on characteristics apparently not related to those proposed by ISO/IEC 25010. The form also contains tables for measures, requirements and device limitations, which might be important to relate, characterize and interpret the quality characteristics present in each paper.

Apart from these fields, the form contains other information such as the authors, the year of publication and the search engines, in order to help with the writing of the thesis. Additional information was also extracted: the software quality standard, if some was used, and the definition of mobile application in the given context, if present.

#### **3.5.2 RQ1: Quality Characteristics from ISO/IEC 25010**

The series of international standards ISO/IEC 25000 identifies eight product quality characteristics and link them to 30 sub-characteristics. It also identifies five quality in use characteristics and associate them with nine sub-characteristics (ISO/IEC, 2011).

A process was followed in order to identify characteristics in the included papers:

- (i) Identify quality characteristics and sub-characteristics, both in use or product quality, associated with mobile applications, where authors explicitly reference ISO/IEC 25000.
- (ii) Identify quality characteristics and sub-characteristics, both in use or product quality, associated with mobile applications, where authors explicitly reference ISO/IEC 9126 (ISO/IEC, 2001), predecessor of ISO/IEC 25010. Both standards were combined using a comparative table in ISO/IEC 25010's annex A (ISO/IEC, 2011).
- (iii) Identify quality characteristics and sub-characteristics, both in use or product quality, associated with mobile applications, where authors do not explicitly reference SQuaRE standards nor its predecessor ISO/IEC 9126. In this case, there are two possible situations: (1) there is an equivalent and similar definition to those in the standards; or (2) the paper does not contain a definition, but the context permits the association.
- (iv) Identify attributes associated with mobile applications equivalent to characteristics and sub-characteristics, both in use or product quality by similarity. In this case it was considered a synonym (e.g. Functional Suitability, Functionality and Functional Quality), following the proposal of (Marinho & Resende, 2012).
- (v) Still following the proposal of (Marinho & Resende, 2012), characteristics and sub-characteristics were added when the paper defines problems or restrictions to the use, suggesting the necessity of a characteristic or sub-characteristic (e.g. "abandoning the use of the application due to faults" suggests the necessity of the characteristic Reliability).

In order to achieve the results, described in Table 3.10 and Table 3.11, the identification was initially performed by a researcher and revised by the same researcher assisted by another researcher. The characteristics and sub-characteristics present in the tables are the same from the models in ISO/IEC 25010.

Table 3.10 – Occurrences of Product Quality model’s characteristics and sub-characteristics.

Characteristics & Sub-Characteristics		Occurrences
<b>Functional Suitability</b>		[M1] [M2] [M5] [M7] [M11] [M16] [M17] [M19] [M21] [M22] [M29] [M31] [M38] [M43] [M51] <i>Subtotal: 15</i>
<b>Sub-characteristics</b>	<b>Functional completeness</b>	[M22] [M29] <i>Subtotal: 2</i>
	<b>Functional correctness</b>	[M8] [M17] [M19] [M22] [M29] [M31] <i>Subtotal: 6</i>
	<b>Functional appropriateness</b>	[M22] [M29] [M38] [M51] <i>Subtotal: 4</i>
<b>Performance Efficiency</b>		[M1] [M2] [M5] [M7] [M11] [M15] [M16] [M18] [M19] [M21] [M22] [M24] [M28] [M29] [M31] [M32] [M38] [M44] [M51] [M43] <i>Subtotal: 20</i>
<b>Sub-characteristics</b>	<b>Time behaviour</b>	[M11] [M22] [M28] [M29] [M30] [M31] [M38] [M44] [M48] <i>Subtotal: 9</i>
	<b>Resource utilization</b>	[M4] [M6] [M11] [M17] [M22] [M28] [M29] [M38] [M44] [M48] <i>Subtotal: 10</i>
	<b>Capacity</b>	[M29] <i>Subtotal: 1</i>
<b>Compatibility</b>		[M1] [M19] [M29] <i>Subtotal: 3</i>
<b>Sub-characteristics</b>	<b>Co-existence</b>	[M22] [M29] <i>Subtotal: 2</i>
	<b>Interoperability</b>	[M22] [M29] [M47] [M48] [M53] <i>Subtotal: 5</i>
<b>Usability</b>		[M1] [M2] [M4] [M5] [M6] [M7] [M8] [M11] [M16] [M17] [M18] [M19] [M21] [M22] [M24] [M29] [M31] [M32] [M37] [M38] [M43] [M45] [M48] [M50] [M51] <i>Subtotal: 25</i>
<b>Sub-characteristics</b>	<b>Appropriateness Recognisability</b>	[M11] [M19] [M22] [M29] [M31] [M38] [M45] <i>Subtotal: 7</i>
	<b>Learnability</b>	[M7] [M11] [M19] [M22] [M27] [M29] [M38] [M41] [M45] [M50] [M52] <i>Subtotal: 11</i>
	<b>Operability</b>	[M7] [M11] [M13] [M17] [M19] [M22] [M29] [M31] [M38] [M42] [M45] [M49] [M50] [M52] [M53] <i>Subtotal: 15</i>
	<b>User error protection</b>	[M19] [M31] [M37] [M41] [M50] <i>Subtotal: 5</i>
	<b>User interface aesthetics</b>	[M11] [M13] [M22] [M29] [M37] [M45] [M52] <i>Subtotal: 7</i>
	<b>Accessibility</b>	[M7] [M11] [M29] [M39] [M37] [M52] <i>Subtotal: 6</i>
<b>Reliability</b>		[M1] [M2] [M3] [M4] [M5] [M8] [M11] [M19] [M16] [M29] [M31] [M43] [M21] [M22] [M24] <i>Subtotal: 15</i>

Characteristics & Sub-Characteristics		Occurrences
Sub-characteristics	<b>Maturity</b>	[M22] [M29] [M31] <i>Subtotal: 3</i>
	<b>Availability</b>	[M3] [M11] [M31] [M35] <i>Subtotal: 4</i>
	<b>Fault tolerance</b>	[M3] [M8] [M9] [M11] [M22] <i>Subtotal: 5</i>
	<b>Recoverability</b>	[M8] [M11] [M22] <i>Subtotal: 3</i>
<b>Security</b>		[M1] [M7] [M8] [M16] [M17] [M19] [M20] [M22] [M29] [M31] [M32] <i>Subtotal: 11</i>
Sub-characteristics	<b>Confidentiality</b>	[M20] [M29] [M31] [M33] [M37] [M38] [M51] <i>Subtotal: 7</i>
	<b>Integrity</b>	[M20] [M29] [M31] <i>Subtotal: 3</i>
	<b>Non-repudiation</b>	[M20] [M29] <i>Subtotal: 2</i>
	<b>Accountability</b>	[M20] [M29] [M32] <i>Subtotal: 3</i>
	<b>Authenticity</b>	[M17] [M20] [M29] [M31] <i>Subtotal: 4</i>
<b>Maintainability</b>		[M1] [M16] [M19] [M22] [M38] <i>Subtotal: 5</i>
Sub-characteristics	<b>Modularity</b>	<i>Subtotal: 0</i>
	<b>Reusability</b>	<i>Subtotal: 0</i>
	<b>Modifiability</b>	[M6] [M22] [M38] [M51] <i>Subtotal: 4</i>
	<b>Analysability</b>	[M22] <i>Subtotal: 1</i>
	<b>Testability</b>	[M37] [M48] <i>Subtotal: 2</i>
<b>Portability</b>		[M1] [M6] [M7] [M13] [M15] [M16] [M17] [M18] [M20] [M21] [M22] [M38] [M48] [M51] <i>Subtotal: 14</i>
Sub-characteristics	<b>Adaptability</b>	[M1] [M13] [M6] [M7] [M15] [M20] [M38] [M51] <i>Subtotal: 8</i>
	<b>Installability</b>	[M20] [M22] [M48] [M52] <i>Subtotal: 4</i>
	<b>Replaceability</b>	[M20] [M22] [M51] <i>Subtotal: 3</i>

Table 3.11 – Occurrences of Quality in Use model's characteristics and sub-characteristics

Characteristics & Sub-Characteristics		Occurrences
<b>Effectiveness</b>		[M9] [M14] [M19] [M23] [M29] [M27] [M41] [M50] [M52] <i>Subtotal: 9</i>
<b>Efficiency</b>		[M9] [M13] [M14] [M16] [M19] [M23] [M27] [M29] [M41] [M42] [M50] [M51] [M52] <i>Subtotal: 13</i>
<b>Satisfaction</b>		[M14] [M19] [M23] [M27] [M29] [M31] [M35] [M41] [M50] [M52]

Characteristics & Sub-Characteristics		Occurrences
		<i>Subtotal: 10</i>
Sub-characteristics	Usefulness	[M19] [M29] [M37] [M42] [M49] <i>Subtotal: 5</i>
	Trust	[M27] [M29] [M31] <i>Subtotal: 3</i>
	Pleasure	[M19] [M27] [M29] [M40] [M42] [M49] <i>Subtotal: 6</i>
	Comfort	[M16] [M19] [M27] [M29] <i>Subtotal: 4</i>
Freedom from risk		[M14] [M19] [M27] [M29] <i>Subtotal: 4</i>
Sub-characteristics	Economic risk mitigation	[M19] [M29] <i>Subtotal: 2</i>
	Health and safety risk mitigation	[M29] <i>Subtotal: 1</i>
	Environmental risk mitigation	<i>Subtotal: 0</i>
Context coverage		[M14] [M16] [M19] [M28] [M29] [M35] [M37] [M41] [M53] <i>Subtotal: 9</i>
Sub-characteristics	Context completeness	[M14] [M16] [M19] [M28] [M29] [M35] [M37] [M41] [M53] <i>Subtotal: 9</i>
	Flexibility	[M6] [M14] [M16] [M19] [M28] [M29] [M35] [M37] [M41] [M53] <i>Subtotal: 10</i>

### 3.5.3 RQ2: Quality Characteristics not present in ISO/IEC 25010

This question aims to identify other quality attributes, which authors identified as necessary for mobile applications and which are not present in ISO/IEC 25010.

We tried to identify new quality attributes in the included papers. Initially 90 attributes were identified. Similarities were observed in an in-depth analysis of them. A final list was drawn-up following the proposal of (Marinho & Resende, 2012).

- (i) Identify attributes listed as important in the included papers, which are not synonyms with ISO/IEC 25010 quality characteristics.
- (ii) Group synonyms or similar attributes according to their definitions.
- (iii) Add an attribute when the included paper defines problems or restrictions to the use, suggesting the necessity of a quality attribute.
- (iv) Select the most suitable definition for the attribute. A definition must be manually composed if the attribute is explained but not defined in the papers.

In order to achieve the results, described in Table 3.12 and Table 3.13, the identification was initially performed by a researcher and revised by the same researcher assisted by another researcher.

Notice, given the results in Table 3.10, Table 3.11, Table 3.12 and Table 3.13, that the papers M12, M25, M26, M34, M36 and M46 did not contribute with occurrences of characteristics nor sub-characteristics. However, considering that they were read and interpreted in search of possible occurrences, they were not excluded from the list of included papers.

Table 3.12 – Product quality characteristics not present in ISO/IEC 25010.

Characteristics & Sub-Characteristics		Definition	Occurrences
<b>Information Quality</b>		The degree to which the mobile application delivers accurate and suitable information, which meets stated and implied needs when used under specified conditions.	[M19] [M31] <i>Subtotal: 2</i>
<b>Sub-characteristics</b>	<b>Information Correctness</b>	The degree to which the mobile application delivers both semantically and syntactically correct information for a given language.	[M31] <i>Subtotal: 1</i>
	<b>Information Credibility</b>	The degree to which the mobile application delivers reputable, objective, and verifiable information.	[M31] <i>Subtotal: 1</i>
	<b>Information Conciseness</b>	Degree to which the information coverage is compact in the mobile application, without being overwhelming.	[M19] <i>Subtotal: 1</i>
<b>Usability (ISO/IEC 25010)</b>		-	-
<b>Sub-characteristics</b>	<b>Navigation</b>	The degree to which the mobile application enables users to find easily the functionality or information they need.	[M19] [M40] <i>Subtotal: 2</i>
	<b>Interface Visibility</b>	The degree to which the mobile application makes appropriate usage and placement of text format that impact positively the user speed of comprehension.	[M19] <i>Subtotal: 1</i>
<b>Use of Clear Forms</b>		The degree to which the use of forms in the mobile application is clear and contains context-sensitive help.	[M19] <i>Subtotal: 1</i>
<b>Use of Minimized Forms</b>		The degree to which forms in the mobile application are minimized.	[M4] <i>Subtotal: 1</i>
<b>Use of Hierarchical Menus</b>		The degree to which the mobile application's menus are limited, simple and easily navigated with a clear breadcrumb path showing where the user has come from and where they can go to.	[M19] [M40] <i>Subtotal: 2</i>
<b>Data Persistence</b>		The degree to which the mobile application keeps information even after it is paused or killed.	[M4] [M6] [M15] [M51] <i>Subtotal: 4</i>

Table 3.13 – Quality in use characteristics not present in ISO/IEC 25010.

Characteristics & Sub-Characteristics		Definition	Occurrences
Sense of Community		The degree to which a mobile application user is satisfied when meeting, collaborating and communicating with other users with similar interest and needs.	[M10] [M19] Subtotal: 2
Usability in Use		The degree to which specified mobile application users can achieve specified goals with effectiveness, efficiency, learnability in use, and without communicability breakdowns in a specified context of use.	[M19] Subtotal: 1
Sub-characteristics	Learnability in Use	The degree to which specified mobile application users can learn efficiently and effectively while achieving specified goals in a specified context of use.	[M19] [M41] [M53] Subtotal: 3
	Continuous Communication	The degree to which specified mobile application users can achieve specified goals without communicative breakdowns in the interaction in a specified context of use.	[M19] Subtotal: 1
	Memorability	The degree to which a mobile application user can effectively retain how to use an application.	[M41] Subtotal: 1

### 3.5.4 Additional Results

The extraction form (Annex A) also includes some extra fields that assisted the achieving of additional results: the definition of mobile application if present, the software quality standard if some was used, the search engines where the paper was found and the year of publication.

#### 3.5.4.1 Definition of Mobile Application

The extraction of definitions intended to identify what a mobile application is from the point of view of the authors. In general, the papers implicitly consider that mobile applications are systems running on smartphones and tablets. Some authors do not even cite these devices nor definitions, and act as if the context was already known, or evident. Most of the papers providing definitions are about specific app categories. Only a few articles provide general definitions. Table 3.14 and Table 3.15 list the specific and general definitions, respectively.

Table 3.14 – Definitions of specific mobile application categories

ID	App Category	Definition
M1	Thick/Thin client apps	Thick client is a kind of application that has many offline data processing which don't require communication with server, while thin client depends heavily on server for data processing.

ID	App Category	Definition
M2	Mobile Personal Health Records	MPHRs are mobile applications that allow the users to record and browse their personal medical information regardless of the location and/or the time.
M4	Data-intensive apps	A data-intensive app is an app that predominantly passes data between one or more APIs and renders that information to the screen.
M5	Mobile Personal Health Records	Mobile personal health records (mPHRs) are mobile applications that allow users to access and record their medical information in any place and at any time by using their smartphones.
M7	Mobile Learning	The main goal is to provide greater motivation, convenience and flexibility to the learning processes in general.
M10	Socially Aware mobile apps	As mobile applications become more pervasive, there is an increasing need for them to exhibit awareness of the social context of the user. A software system or application is socially aware, if it uses social context information such as social roles, relationships, interactions and situations, to adapt its behavior.
M11	Mobile Personal Health Records	In order to improve the management of the patients' health data and promote the exchange between the patients and healthcare providers, mobile personal health records, as mobile applications, are used to access, store and manage these data.
M13	Mobile travel apps	Provides various information including destination information, hotel recommendations, local customs and culture, as well as local delicacies. Thus, travelers can plan in advance with the mobile travel apps. On the other hand, mobile travel apps provide travel related companies a direct channel to create and maintain the conversation with customers before, during, and after a trip.
M16	Mobile business apps	An application that is integrated to an existing IT infrastructure and runs on a mobile device like a smartphone or tablet within a business environment.
M17	Mobile Learning	The ability of using handheld devices to access learning resources.
M28	Context-aware apps	Context-aware mobile applications are even more complex than their non-mobile distributed counterparts, since they involve connections between executable software components or objects that can migrate from node to node within a heterogeneous software and networking environments.
M36	Content-oriented apps	Fulfill individual needs for information, entertainment, communication, productivity and socialization.
M36	Marketing-oriented apps	Are mostly used by companies for brand advertising or promotion.
M36	Service-oriented apps	Let users perform tasks - for example, check a train schedule, book theater tickets, or shop at a mobile commerce platform.
M42	App-based mobile tour guide	AMTG is defined as a mobile device-installed travel app that provides location-dependent guidance and engaging interaction when a traveler arrives at a destination.
M46	Mobile business app	Mobile business applications are usually tailored to a mobile device such as a smartphone or a tablet (not a laptop), integrated into an existing IT infrastructure, are task-oriented and focused on a clear and limited scope of functionality, and are based on the mobility potential of a company's business processes.



Table 3.15 – Definitions of mobile application

ID	Definition
M6	When we talk about mobile software in this paper, we restrict ourselves to software of current and future mobile devices like mobile phones, tablets based on mobile platforms and other interactive and restricted embedded mobile hardware.
M18	The smart mobile devices, like smartphones and tablets, are becoming pervasive. These devices are characterized by a wide range of interaction possibilities and some restrictions, which are not usually considered for non-portable computers.
M19	WebApps, a combination of information, integrated functionalities and services have become the most predominant form of software delivery today with users and businesses choosing to rent or use software rather than buy it.
M22	Traditional applications use primarily in front of a computer, and mobile applications are used anywhere with hand. Traditional primarily with the mouse, and mobile applications primarily through the fingers to complete the operation.
M25	Handset terminals have experienced a shift from being simple communication devices to become high-end, multipurpose computer equipment. Smartphones are driven by powerful operating systems that allow users to add and remove applications, and they employ architecture that is similar to a regular personal computer.
M26	State-of-the-art mobile devices, with their mass of sensors, have the ability to identify plenty of contexts by which we are surrounded.
M27	Mobile devices and their applications provide significant advantages to their users, in terms of portability, location awareness, and accessibility.
M32	Platform fragmentation, the physical characteristics of the mobile device, user experience, integration of third-party apps, performance, security, system integration, and the deployment of mobile apps are pressing concerns that need to be addressed.
M37	These applications are developed to run on mobile devices, and to allow user mobility.
M40	Mobile devices play an important role in the modern society. They are being used by people of all social groups for various purposes. They can be found in the fields of education, entertainment, medicine, communication service, military systems, and so on.
M43	Mobile apps are software applications developed for use on mobile devices such as smartphones and tablets. Once developed, an app is sold via an application distribution platform, commonly known as an app store.
M53	Mobile devices accompany their users throughout much if not all of the day. Unlike stationary work or home computers, or even laptops that are taken to specific places such as meetings and airports, mobile phones are with us in all the indoor and outdoor environments we travel.

#### 3.5.4.2 *Software Quality Standards*

The preliminary execution of the search string returned plenty of papers about the quality of mobile applications basing their results in ISO/IEC 9126 or ISO/IEC 25010. Given the degree of formality found in the literature, from this point the research questions were updated to separate the quality characteristics by origin, as discussed in sections 3.5.2 and 3.5.3. A field to store the software quality standard was also included in the extraction form. In total 25 papers used a standard, nearly half of them. Figure 3.2 contains a pie chart showing the proportions.

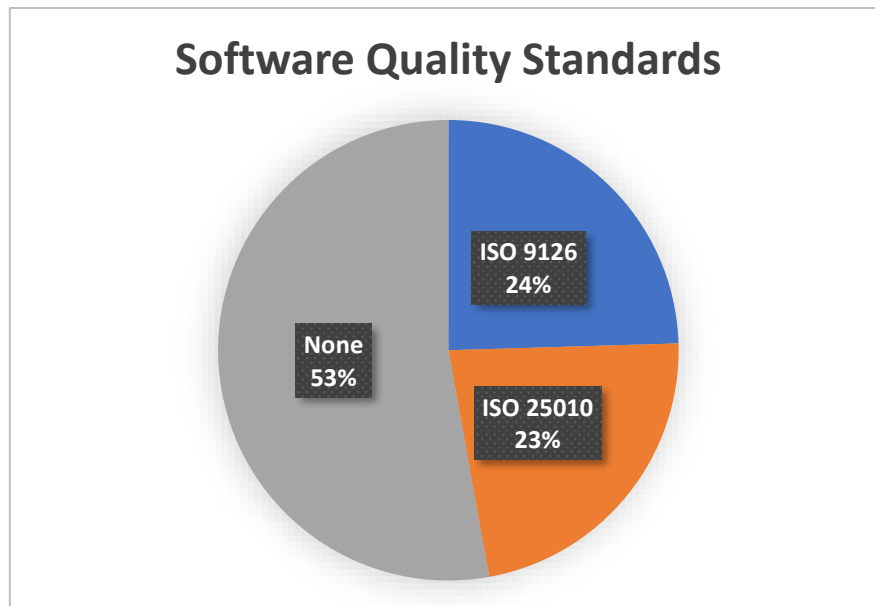


Figure 3.2 – Percentage of use of each software quality standard.

#### 3.5.4.3 *Presence in Search Engines*

At all, five search engines were selected for the execution of the search string. Table 3.7 and Table 3.8 shows the occurrence of papers in each of them. Out of curiosity, the ones included via snowballing were also manually searched in each engine. A closer look at these tables reveal that almost every paper was found in Scopus. For those included by the execution of the search string, only six were found elsewhere, and for the snowballing papers, six were not found in Scopus nor in any of the other engines and one was only found in Web of Science. It represents a coverage of 75.5% of the results by Scopus.

Scopus indexes the content of other engines. The titles of M26, M27, M28, M33, M34 and M35, the papers originally not found by the execution of the string, were directly searched and found in Scopus. The reason for not being returned before lies on the different indexing of the papers in each engine. A better calibration of the search string could possibly have returned all the included papers in one go.

#### 3.5.4.4 *Publication Frequency*

The academic interest in a topic may increase or decrease over the years. Analyzing the frequency of publications regarding the quality of mobile applications may help us to identify if it is an emerging or abandoned approach. Figure 3.3 contains a bar chart relating years and number of papers. Only a few papers are earlier than 2009. Android and iOS technologies emerged in 2007 and popularized the distribution model

of the app stores (Cortimiglia et al., 2011). From this point onwards, it became easier to develop and download applications. It may be the reason for the sudden interest in this research area after 2009, which remains relatively unchanged ever since.

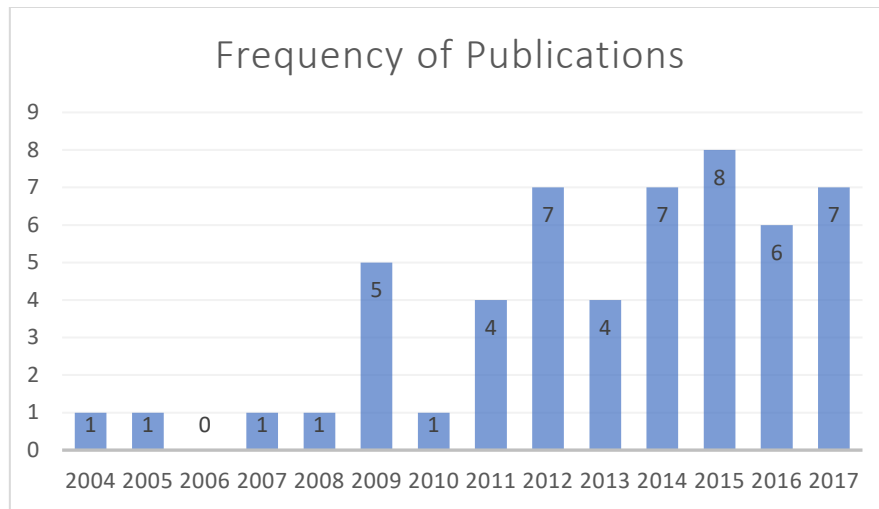


Figure 3.3 – Frequency of publications over the years

### 3.6 Discussion of the Results

The selected papers explore a range of elements: quality of specific categories of mobile applications, researches into mobile device limitations, researches into the importance of a single quality characteristic, proposal of quality models, guidelines, measures, requirements, among others. This variety positively influenced this study and diversified the conclusions about the pertinence of quality characteristics in the context of mobile applications.

The results in section 3.5.2 indicate that usability is the most addressed product quality characteristic, followed by performance efficiency, functional suitability and reliability. As for quality in use, efficiency, satisfaction and context coverage are the most addressed quality characteristics. Regarding the sub-characteristics of both quality models, the most frequent ones are consequently related to usability, performance efficiency and satisfaction. The results in section 3.5.3 indicate the importance of additional quality characteristics strongly related to usability, thereby reinforcing the results obtained from the previous question. Some of the proposed new quality characteristics are information quality, data persistence and sense of community.

Some papers briefly discuss inherent characteristics and limitations of mobile devices and their impact on the quality of the applications. Common limitations such as

battery consumption, CPU usage and low memory directly affect the performance efficiency. Limited screen size, few resolutions and limited input directly affect usability. A frequently mentioned limitation concerns the instability of internet connections, due to the mobility of devices, which restricts the use of wireless network. It might outline the importance of reliability and context coverage. Other frequently mentioned limitation is the occurrence of unexpected interruptions (e.g. receiving a call), which outlines the importance of the proposed quality characteristic data persistence. The characteristic sense of community is a concern related to the increasing popularity of the socialization nature of mobile applications.

The processes defined in sections 3.5.2 and 3.5.3 are strongly based on the interpretation of synonyms. Despite the existence of ISO/IEC 25010 and ISO/IEC 9126, quite a lot of papers used their own names and definitions for referring to a quality characteristic. The list of synonyms became an output of this study (appendix), as part of interpreting the answers for RQ1 and RQ2. It may assist future researches in assembling search strings consisting of quality characteristics.

### **3.7 Threats to Validity**

This section presents threats to validity identified in this research work, together with attempts to reduce them. They should be dealt with carefully not to compromise the validity of the results. The classification presented by (Petersen et al., 2015) will be followed, which considers the importance of descriptive validity, theoretical validity, interpretative validity and generalizability.

Descriptive validity is the extent to which observations are described accurately and objectively. The extraction form, described in section 3.5.1, objectively kept relevant information present in each paper regarding the research questions. However, the extraction may not guarantee the correctness of the results. In general, a quality characteristic was considered and included in the extraction form as long as the paper presented a description and an adequate reason for considering it pertinent, in spite of not presenting a robust validation for their results, in some cases. Furthermore, some papers provided quality characteristics with differing definitions from ISO/IEC 25010, which demanded a deep interpretation of what the authors meant.

Theoretical validity is determined by our ability to be able to capture what we intend to capture. The search string may not have captured many other relevant papers.

The selected control group papers were frequently returned during the evolution of the search string, but they were not returned by every search engine, which might indicate the necessity of further refinement. In addition, the interpretation of the abstracts and the manual selection of papers via snowballing may have biased the results. In order to reduce the bias, both the inclusion and extraction of papers were reviewed several times by a second researcher.

Interpretive validity is achieved when the conclusions drawn are reasonable given the data, and hence maps to conclusion validity. The research questions RQ1 and RQ2 were subjective and demanded effort to both understand what to be extracted and how to interpret the extracted data. To reduce bias, the results were debated over several meetings until a consensus was reached.

Finally, considering the generalizability, the quantity of papers may have been relatively small, possibly due to a too restrictive search string. However, the search engines provided good coverage of the topic, especially Scopus.

### **3.8 Conclusion**

This chapter presented the findings of a systematic mapping, which identified pertinent quality characteristics in the context of mobile applications. We selected 53 papers among the results of the execution of a search string in five search engines. They were extracted and analyzed against two research questions. The results indicated which quality characteristics from the quality models in ISO/IEC 25010 are more pertinent to the context of mobile applications. It also identified 15 quality characteristics not covered by ISO/IEC 25010.

The next chapter presents the development of a survey, aiming at the opinion of mobile users, intended to confirm the results of the systematic mapping.

## 4 Survey

This chapter presents the development and the execution of a survey, intended to confirm the results of the systematic mapping related to the pertinence of quality sub-characteristics in the context of mobile applications. Section 4.1 describes the goal and the lists the sub-characteristics present in the survey. Section 4.2 describes the instrument used in the survey. Section 4.3 presents the execution of the survey. Section 4.4 presents the data analysis. Section 4.5 presents the discussion. Section 4.6 presents the threats to validity. Finally, section 4.7 presents the conclusion.

### 4.1 Goal and Sub-Characteristics

A survey with mobile users was conducted to confirm some of the results of the systematic mapping described in chapter 3. The systematic mapping enabled the ranking of the quality sub-characteristics based on occurrences in studies, yet it is not feasible to compile a final list of essential quality sub-characteristics for the context of mobile applications only with this information.

Literature papers were analyzed in search of occurrences of ISO/IEC 25010 quality characteristics and sub-characteristics. Also, in search of occurrences of additional specific characteristics of this type of application. The number of occurrences was used as a deciding factor in order to define if a certain characteristic would be considered essential or not.

Due to the generality of characteristics, only sub-characteristics were taken into consideration in the final list of results. Three intervals were defined based on the range of occurrences of sub-characteristics, which were comprised between zero and fifteen. The sub-characteristics with less than five occurrences were dismissed (Table 4.2), and those with eight or more occurrences were considered essential (Table 4.3). The remaining ones, with an intermediary number of occurrences (five to seven), were added to the survey for further investigation (Table 4.4).

The systematic mapping also identified 15 quality attributes not previewed by ISO/IEC 25010 (Table 3.12). They will be referred as "attributes" to distinguish from the quality sub-characteristics in ISO/IEC 25010. They were directly added to the survey regardless of their number of occurrences. To reduce the number of questions in the survey, some attributes were adapted. **Use of Clear Forms** and **Use of Minimized Forms**

were merged. **Usability in Use** and **Learnability in Use** were suppressed. **Information Quality** was directly included instead of their sub attributes **Information Correctness**, **Credibility** and **Conciseness**. The resulting list is summarized in Table 4.1. At all, ten quality sub-characteristics and nine quality attributes were added to the survey, totaling 19 questions.

Table 4.1 – Quality attributes, not present in ISO/IEC 25010 quality models

Quality Attribute	Occurrences in the Mapping
Data Persistence	4
Information Quality	2
Use of Hierarchical Menus	2
Sense of Community	2
Navigation	1
Interface Visibility	1
Use of Clear and Minimized Forms	1
Continuous Communication	1
Memorability	1

Table 4.2 – Quality sub-characteristics not considered essential in the context of apps (< 5)

Quality Model	Quality Characteristics	Quality Sub-Characteristics	Occurrences in the Mapping
Product Quality	Functional Suitability	Functional Completeness	2
		Functional Appropriateness	4
	Performance Efficiency	Capacity	1
	Compatibility	Coexistence	2
	Reliability	Maturity	3
		Availability	4
		Recoverability	3
	Security	Integrity	3
		Non-Repudiation	2
		Accountability	3
		Authenticity	4
	Maintainability	Modularity	0
		Reusability	0
		Analysability	1
		Modifiability	4
		Testability	2
	Portability	Instalability	4
		Replaceability	3
Quality in Use	Satisfaction	Trust	3
		Comfort	4
	Freedom From Risk	Economic Risk Mitigation	2
		Health Risk Mitigation	1
		Environmental Risk Mitigation	0

Table 4.3 – Quality sub-characteristics considered essential in the context of apps ( $\geq 8$ )

Quality Model	Quality Characteristics	Quality Sub-Characteristics	Occurrences in the Mapping
Product Quality	Performance Efficiency	Time Behavior	<b>9</b>
		Resources Utilization	<b>10</b>
	Usability	Learnability	<b>11</b>
		Operability	<b>15</b>
	Portability	Adaptability	<b>8</b>
Quality in Use	Effectiveness	Effectiveness	<b>9</b>
	Efficiency	Efficiency	<b>13</b>
	Context Coverage	Context Completeness	<b>9</b>
		Flexibility	<b>10</b>

Table 4.4 – Quality sub-characteristics added to the survey for further investigation (5 - 7)

Quality Model	Quality Characteristics	Quality Sub-Characteristics	Occurrences in the Mapping
Product Quality	Functional Suitability	Functional Correctness	<b>6</b>
	Compatibility	Interoperability	<b>5</b>
	Usability	Appropriate Recognisability	<b>7</b>
		User Error Protection	<b>5</b>
		User Interface Aesthetics	<b>7</b>
		Accessibility	<b>6</b>
	Reliability	Fault Tolerance	<b>5</b>
	Security	Confidentiality	<b>7</b>
Quality in Use	Satisfaction	Usefulness	<b>5</b>
		Pleasure	<b>6</b>

## 4.2 Instrument

The tool selected to host the survey was LimeSurvey<sup>10</sup>, an open source tool to conduct online surveys. LimeSurvey has a responsive interface, which properly fits mobile device resolutions. It was a deciding factor, as we assumed that respondents would prefer to answer from their mobile phones instead of using a computer.

The respondents should opine over the importance of a set of 19 quality characteristics in the context of mobile applications. The definitions of quality characteristics are technical and might be hard to understand for people not related to Software Engineering. Including these definitions as questions would increase answer time and derail the participation of people from other fields of study. The elaboration of the questions' titles was tricky because they could not be neither the name of the

<sup>10</sup> <https://www.limesurvey.org/>



characteristic nor its official definition. The simpler it was, the more answers the survey would obtain, so the chosen strategy was to develop a small and informal question based on the definition. Table 4.5 and Table 4.6 list the questions developed for each one of the quality characteristics identified in the systematic mapping. Even though these tables present the questions in English, the original instrument was made entirely in Portuguese.

The survey was composed of two parts. The first part collected demographic data: the **gender**, the **age** group and the **level of education**. The second part collected substantive data. The first question inquired about the respondent's favorite app category, given five possibilities: **mobility** apps (e.g. Uber), **food delivery** apps (e.g. iFood), **tourism** apps (e.g. Trivago), **messaging** app (e.g. WhatsApp) and **banking** apps (e.g. Nubank). The respondents were guided to answer questions about the quality characteristics considering only the specific selected app category.

Table 4.5 – Survey questions for the ISO/IEC 25010 quality sub-characteristics

Quality Sub-characteristics	Survey Question
Confidentiality	How much do you think that is important for the app to never use your personal information without your authorization?
User interface aesthetics	How much do you think that is important for the app to have a beautiful appearance?
Appropriateness Recognisability	How much do you think that is important for the users, when using the app for the first time, perceiving if it is appropriate for their needs?
Accessibility	How much do you think that is important for the app to have adaptations to permit its use by people with hearing, visual and motor limitations?
Functional Correctness	How much do you think that is important for the app to correctly do what it is expected of it?
Pleasure	How much do you think that is important for the app's use to be pleasant?
Interoperability	How much do you think that is important for the app to communicate data with other apps?
User Error Protection	How much do you think that is important for the app to avoid that the users commit mistakes?
Fault Tolerance	How much do you think that is important for the app to behave properly even when there are problems with the software and the device?
Usefulness	How much do you think that is important for the app to be useful for the users in order to help them achieving their needs?

Table 4.6 – Survey questions for the quality attributes not previewed by ISO/IEC 25010

Quality Sub-characteristics	Survey Question
Information Quality	How much do you think that is important for the app to provide accurate and adequate information for the user's needs?
Navigation	How much do you think that is important for the app's information and functionalities to be easily found within the application?
Interface Visibility	How much do you think that is important for the app to have well-organized screens in order to facilitate the understanding?
Use of Clear and Minimized Forms	In case of apps with forms, how much do you think that is important for them to be clear and with help?
Use of Hierarchical Menus	How much do you think that is important for the app to have few, simple and easy to navigate menus?
Data Persistence	Sometimes we have to take a phone call while using as app. When the call ends, how much do you think that is important that the app keeps the appropriate information when the app was paused?
Sense of Community	How much do you think that is important for the app to facilitate social interaction?
Continuous Communication	How much do you think that is important that problems like network disconnection and low GPS signal do not disturb the app's usage?
Memorability	We usually forget how to use the app after a time not using it. How much do you think that is important for the app to be easily remembered after some time not being used?

Each question about quality characteristics contained a title, a VAS (visual analog scale) option (Wewers & Lowe, 1990) and a multiple-choice with two options. The multiple choice' options should have been checked if the respondent did not know how to answer or were not sure if the question was related to the selected app category. Figure 4.1 illustrates the composition of these elements for the sub-characteristic **Confidentiality**.

How much do you think that's important for the app to never use your personal information without your authorization?

Answer

A little A lot

☐ I don't know how to answer the question above.

☐ I don't think that this question is related to the type of app that I selected.

Figure 4.1 – Example of VAS question

The VAS questions store a floating-point value between 0 and 10, visually selected with a slider. This scale is usually used in psychological studies and it allows all

arithmetic calculus (Wewers & Lowe, 1990). The VAS question consists of a horizontal line with two anchor points. Instead of the numbers, we placed the labels "a little" and "a lot." In the end of the survey, a short text indicated that the respondent might answer the survey again for another app category, if they wanted to.

#### 4.2.1 Target Audience

The survey was targeted to anyone above 18 years old and living in Brazil, regardless of the field of work, **education level** or **gender**. We assumed that respondents had at least a little experience as mobile users. The survey was open and people were requested either directly or through social network posts.

#### 4.2.2 Pilot Test

The survey was sent to a selected group of five people as part of a pilot test. These people had different **ages**, **genders** and **education levels**. At this time, the second multiple-choice option below the VAS questions, "I don't think that this question is related to the type of app that I selected", was not yet included.

Only one respondent was male. One respondent had incomplete college degree; all the other respondents had complete college degree. The frequency of the age groups was well-balanced. One respondent answered twice, which was pointed as a possibility in the end of the survey. Every respondent selected the **messaging** apps category, except for the respondent who answered twice, who also selected the **banking** apps category. Their answers in the VAS questions were well-balanced and made sense as a whole.

After answering, they filled a form about the survey's layout and comprehensibility, and with the possibility of improvement suggestions. The reception was positive. One respondent suggested the inclusion of an option for the case of questions not related to the app category, and a new multiple choice was indeed included. Their average responding time was of five minutes.

### 4.3 Execution

The survey request was directly sent by email, Facebook Messenger and WhatsApp to several people; shared in Facebook and LinkedIn and promoted by a Facebook page<sup>11</sup>. It was carried out from June 12, 2019 to June 21, 2019 and we had 500

---

<sup>11</sup> <https://www.facebook.com/business/help/347839548598012>

valid answers. The population was broad, considering that, except by the age restriction, the survey could be answered by any user of mobile applications, regardless of other characterizations. From a statistical point of view, the number of responses may have been a threat to validity.

The survey also had 180 invalid answers. Answers were not considered valid unless the respondents indeed answered the survey until the end, both demographic and substantive questions. Almost every respondent who did not finish the survey completed the characterization questions and ignored the questions about the quality characteristics.

For every answer, we donated R\$1.00 to a Brazilian project which assists children with heart diseases. This initiative not only helped the institution, but also created empathy for the survey and potentially made people answer more than once. A total of R\$500.00 were donated to Pro Criança Cardíaca<sup>12</sup> in Rio de Janeiro.

#### 4.4 Data Analysis

Primarily, we analyzed the results of the descriptive data (characterization) that are showed in Figure 4.2, Figure 4.3 and Figure 4.4. Then, Figure 4.5 shows the results of the first substantive question, about the app category. Given the figures, we can note the following:

- Regarding the **gender** (Figure 4.2), 292 participants (58.4%) of the participants were female and 208 participants (41.6%) were male.

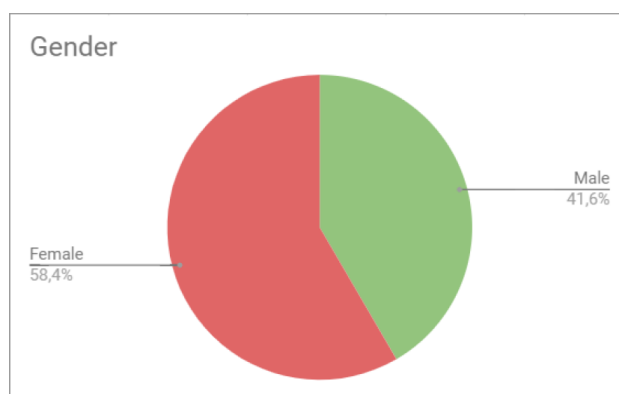


Figure 4.2 – Gender of the participants

- Regarding the **education level** (Figure 4.3), two participants (0.4%) did not finish primary education, eleven participants (2.2%) finished primary

---

<sup>12</sup> <http://www.procrianca.org.br/>

education, eleven participants (2.2%) did not finish high school, 34 (6.8%) participants finished high school, 69 (13.8%) participants did not finish college and 373 (74.6%) participants finished college. Almost all the respondents were graduated, in spite of the fact that the target audience was much broader.

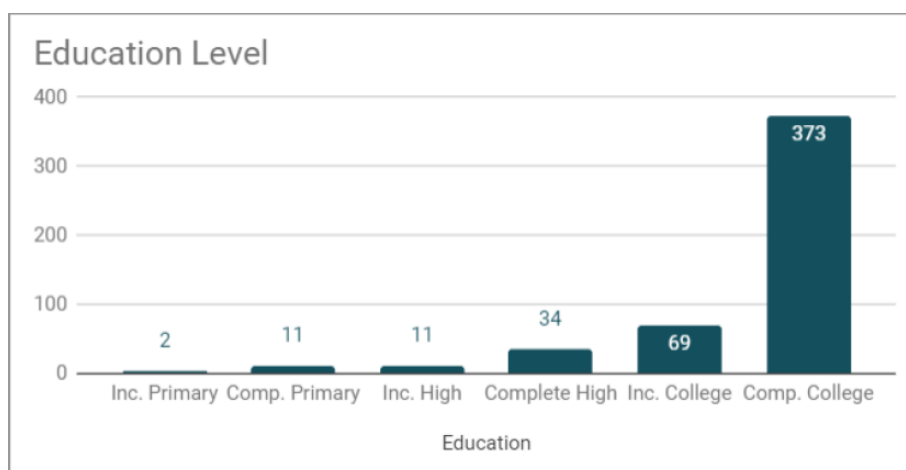


Figure 4.3 – Education Level of the participants

- Regarding the **age** (Figure 4.4), 145 participants (29%) aged between 18 and 30; 139 participants (27.8%) aged between 31 and 40; 73 participants (14.6%) aged between 41 and 50; 87 (17.4%) participants aged between 51 to 60; and 56 participants (11.2%) aged above 60. The number of participants above 60 was high.

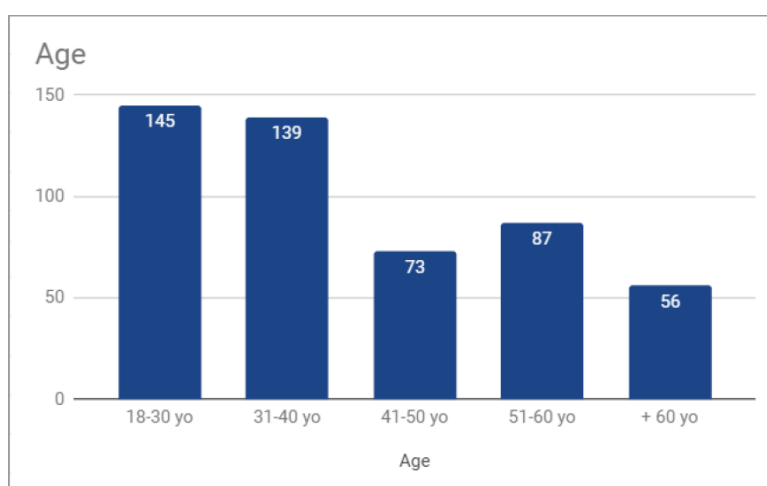


Figure 4.4 – Age of the participants

- Regarding the selected app category (Figure 4.5), 42 participants (8.4%) selected **mobility** apps, 15 (3%) participants selected **food delivery** apps,

eleven participants (2.2%) selected **tourism** apps, 351 participants (70.2%) selected **messaging** apps and 81 participants (16.2%) selected **banking** apps. The preference for **messaging** apps was expressively higher than the other categories, presumably due to the popularity of WhatsApp.

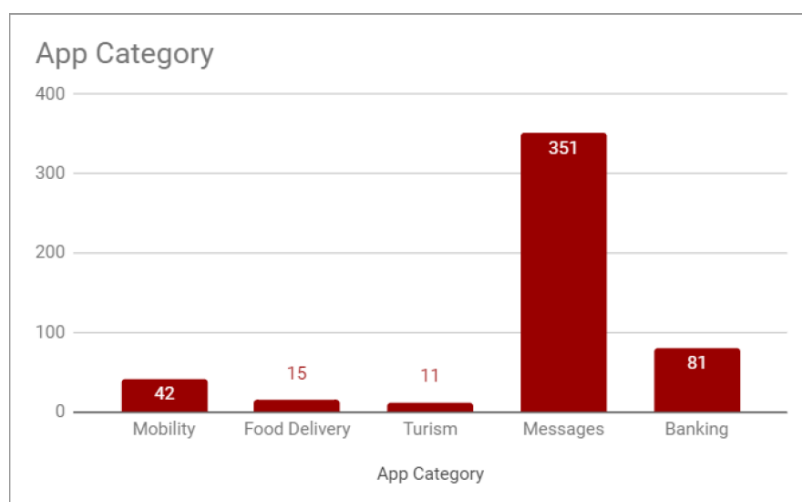


Figure 4.5 – App Categories selected by the participants

Five different results derived from the VAS questions. For each one of them, respondents indirectly selected a value between zero and ten. The answers were grouped by app category and an average was calculated. Some results in Figure 4.6 may be observed and discussed:

- Most of the averages were above 7.0, indicating that the characteristics might be indeed essential to the context of the selected app category.
- Only eleven participants selected **tourism** apps, yet the yellow line in Figure 4.6 is significantly similar to the other lines. The similarity might indicate that individual opinions are not so different from the average itself.
- Independently of the app type or the number of answers, the ISO/IEC 25010 quality sub-characteristic **Interoperability** was not evaluated as important. Its overall average was of 4.73 and might not be considered a essential quality characteristic.
- **Sense of Community**, whose overall average was 5.36, might not be considered essential. It was not evaluated as important for every app

category except the **messaging** app category, which is indeed the one with more elements of social interaction.

- The ISO/IEC 25010 sub-characteristics **User Error Protection** and **Fault Tolerance** had the greatest averages in the **banking** app category, probably due to the seriousness of its functionalities.
- Even though **messaging** apps do not have forms at all, the average for **Use of Clear and Minimized Forms** was considerably high. The lowest average was for the **tourism** app category, which usually have long forms. These results might indicate a misinterpretation of the respondents or some factor that biased the results.
- **User Interface Aesthetics** had the second worst average from the list of ISO/IEC 25010 quality sub-characteristics, while **Functional Correctness** had the best average. It may indicate a preference of functionality over appearance.

## 4.5 Discussion

The survey, which was only announced in Brazil, investigated mobile users' preferences over mobile applications. At all, 500 valid answers were obtained, covering multiple **age** groups and **education levels**.

### 4.5.1 Favorite App Category

The respondents answered about their favorite app category, given a list of five common categories. An expressive quantity of 351 respondents, 70.2% of the total, selected the **messaging** app category. **Banking** apps were the second most selected: 81 respondents, 16.2% of the total. **Food delivery** and **tourism** apps obtained the worst results. **Banking** apps makes it easier to paying bills, a serious activity which people need to execute monthly. This functionality might be useful enough to make respondents forget Uber and iFood, which are also very popular in Brazil. The quality attribute Memorability had its highest average for the **banking** apps, possibly due to the monthly use.

The popularity of apps like WhatsApp and Telegram in Brazil is considerably high, which explain the result and might indicate a cultural bias. WhatsApp is very popular in Brazil; the app doubled the number of users from 2014 to 2019<sup>13</sup>.

---

<sup>13</sup> <https://www.messengerpeople.com/pt-br/whatsapp-no-brasil/>, visited in July 2019

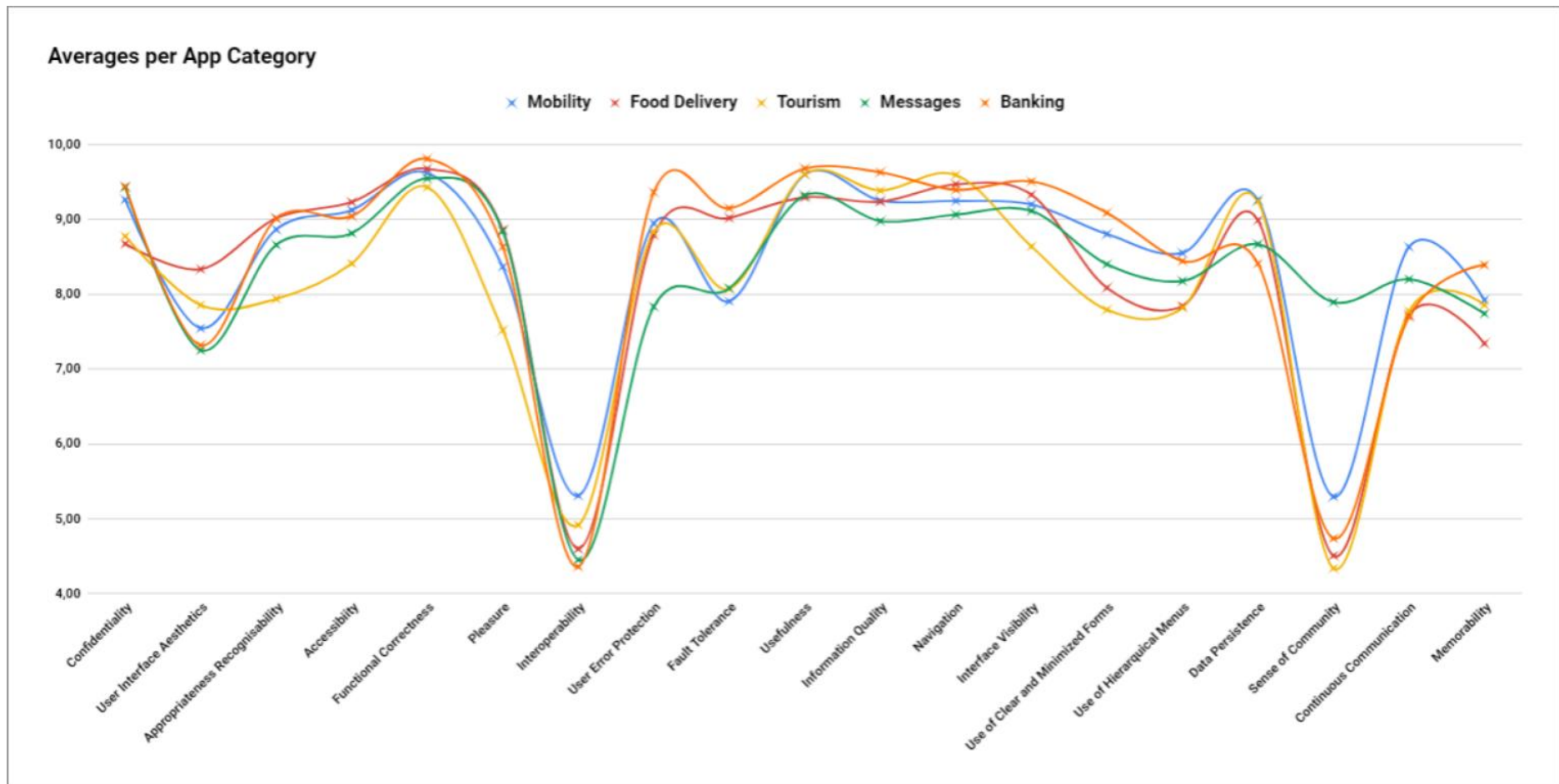


Figure 4.6 – Average of the quality characteristics per app category



#### 4.5.2 Characterization

Four **age** groups were set as options: 18 to 30, 31 to 40, 41 to 50, 51 to 60, and over 60. The quantity of respondents over 51 years old was 143, 28.6%, nearly the same as the 18-30 **age** group. In general, the **ages** of the respondents were well-balanced.

Considering the **gender**, the quantity of female respondents was greater, although the difference in relation to the male respondents was not expressive. Both quantities were nearly the same.

The survey obtained interesting results for the **level of education**. The question considered six possible answers: primary school, high school and college, each of them complete or incomplete. The survey was shared in many channels besides the academic context. People from diverse contexts, jobs and areas of expertise were contacted and directly asked to answer. A Facebook post was also promoted, in which the target audience only considered the location (Brazil) and the age (18 years old onwards). Many people contacted the authors and confirmed that had answered more than once. In spite of this care with the diversity of the audience, only two respondents had incomplete primary school and only eleven had complete primary school. A massive number of 373 respondents, 74.6%, had complete college. The quantity of graduated respondents might indicate a concern with an ongoing study and with academy itself. People from the academic context might have been more aware of the importance of a survey, and consequently spent some time answering it.

However, the **level of education** did not affect the results. The average of every VAS questions altogether, considering all 500 answers was 8.36, while the averages considering only answers from each specific **level of education** were 9.24 for incomplete primary school, 8.71 for complete primary school, 8.79 for incomplete high school, 8.58 for complete high school, 8.24 for incomplete college and 8.36 for complete college.

#### 4.5.3 Answer Time and Answers per day

Both the starting and ending times were stored for each respondent, so it was possible to calculate the average time. Considering characterization and VAS questions, the survey contained 23 questions. The overall average time was of 5:54, an average of 15 seconds per question. It was a good result, which indicates that the survey was indeed very easy and simple. Regarding the number of answers per day, most of the responses happened in the first week, especially in the first Friday. The number of answers decayed

from this Friday on and only increased again in the last two days, due to a final effort of the authors in order to reach 500 responses. By the end of the second week, new respondents were informed about the survey's average answer time when requested to answer. Despite this, people were very resistant and unconcerned. Figure 4.7 shows the number of answers per day.

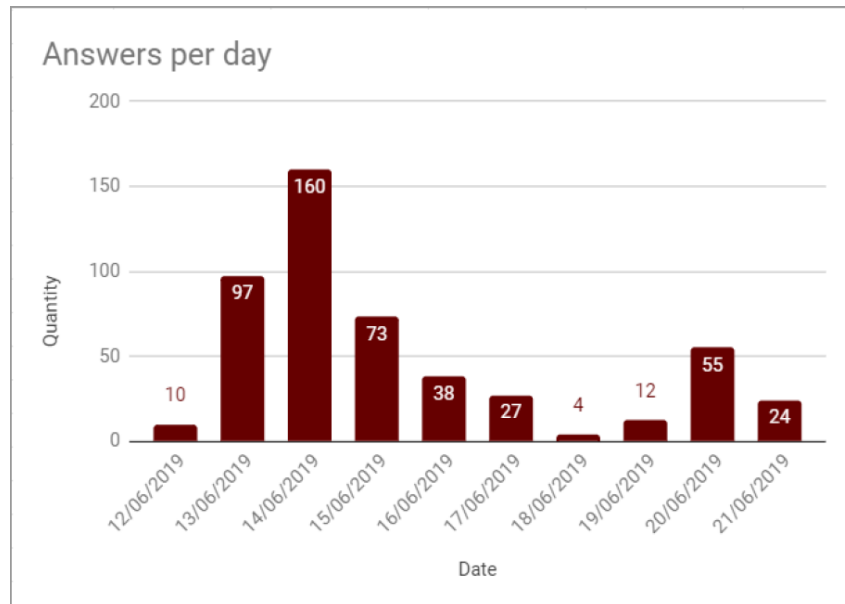


Figure 4.7 – Quantity of answers per day

#### 4.5.4 Multiple Choice Questions

Each VAS question was followed by multiple-choice options, which the respondents should have checked if they were not confident about answering the correspondent VAS question (Figure 4.1).

Checking the first multiple-choice question, “I do not know how to answer the question above,” meant that the respondent did not understand the meaning or the relevance of the VAS question and consequently decided not giving an opinion. Occurrences of this multiple choice might indicate that the quality characteristic is too difficult to understand or that the writing of the VAS question was not made easy enough.

Checking the second multiple-choice question, “I do not think that this question is related to the type of app that I selected,” meant that the respondent properly understood the VAS question and thinks that that topic is not related to the selected app category.

Considering the first multiple-choice question, the characteristic **Appropriateness Recognisability** had the best results, without any checks for any app

category, meaning that all the respondents understood and were able to give an opinion. The second characteristic with the best results was **Pleasure**, which only had three checks in the **mobility** app category. The characteristics with the greatest number of checks were **Fault Tolerance** and **Continuous Communication**, with 25 and 16 occurrences respectively. The VAS questions for these two characteristics were long and probably still too technical.

Considering the second multiple-choice, the characteristic **Functional Correctness** and **Accessibility** had the best results, with only one and two checks for the **messaging** app category, respectively. The characteristics with the greatest number of checks were **Use of Clear and Minimized Forms** and **Sense of Community**. **Use of Clear and Minimized Forms** impressively had 62 checks for the **messaging** app category. This result clashes with the average of this characteristic for the **messaging** app category, which was considerably high. This might mean that the respondents who gave an answer in the VAS question probably misunderstood the question or the meaning of the word “Form”, which might have been interpreted as any kind of data input. Regarding **Sense of Community**, there were 17 checks for the **banking** app category, which is in accordance with the nature of this app category.

Most of the quality characteristics did not obtain any checks for the categories **food delivery** and **tourism**; however, this result should be carefully analyzed, given that the samples for these categories were too small: only fifteen and eleven respondents, respectively.

#### 4.5.5 Quality Characteristics per App Category

In Figure 4.6, all five lines follow a considerably similar path. The results and interpretations are different for each quality characteristic in each app category, but in general, the lines rise and fall together. The red (**food delivery**) and yellow (**tourism**) lines contain very small samples (fifteen and eleven, respectively), however they are similar to the other three lines, with greater samples. This might indicate that the opinion of single respondents is similar to the average.

Regarding **mobility** apps, the least essential characteristics were **Interoperability** and **Sense of Community**. The most essential were **Functional Correctness** and **Usefulness**. **Fault Tolerance** also had a low average when compared to the other categories, and the reason might be the server-side nature of the **mobility** apps category:

the drivers wait for a passenger, and passengers point out that they need a driver. As long as a driver and a passenger find themselves, only the driver's GPS needs to work properly and no other action is required. From this point of view, the reason for Interoperability's low average in this category is not evident.

For **food delivery** apps, the least essential characteristics were **Interoperability** and **Sense of Community**. The most essential were **Functional Correctness** and **Navigation**; **Usefulness**, **Interface Visibility** and **Information Quality** also had high averages. It was also the greatest average of **Accessibility** among all. Although, the large amount of high averages should be analyzed with care, as long as the sample for this category is just composed of 15 respondents.

Regarding **tourism** apps, the least essential characteristics were, again, **Interoperability** and **Sense of Community**. The most essential were **Navigation** and **Usefulness**. Oddly, **Pleasure** had its third lower average. The results for this category should be inferred with care because of the sample size, only composed of eleven respondents.

For **messaging** apps, the least essential characteristics were Interoperability and **User Interface Aesthetics**. The most essential were **Confidentiality** and **Functional Correctness**. The low importance of **User Interface Aesthetics** might mean that the use of these apps happens due to the functionality itself, and users do not care too much about the appearance. Apps for exchanging messages are simple and usually do not demand a complex interface. These apps constantly communicate personal information, so **Confidentiality** is indeed necessary. The attribute **Use of Clear and Minimized Forms** had a strange result: its average it was considerably high, even though **messaging** apps not usually contain forms. This result might indicate that the question for this quality attribute was not clear enough.

Regarding **banking** apps, the least essential characteristics were **Interoperability** and **Sense of Community**. Many characteristics were considered essential: **Functional Correctness**, **Usefulness**, **Information Quality**, **Interface Visibility** and **Confidentiality**. In addition, **Fault Tolerance** and **User Error Protection** had the greatest averages among all app categories, probably due to the seriousness of the performed functionalities.

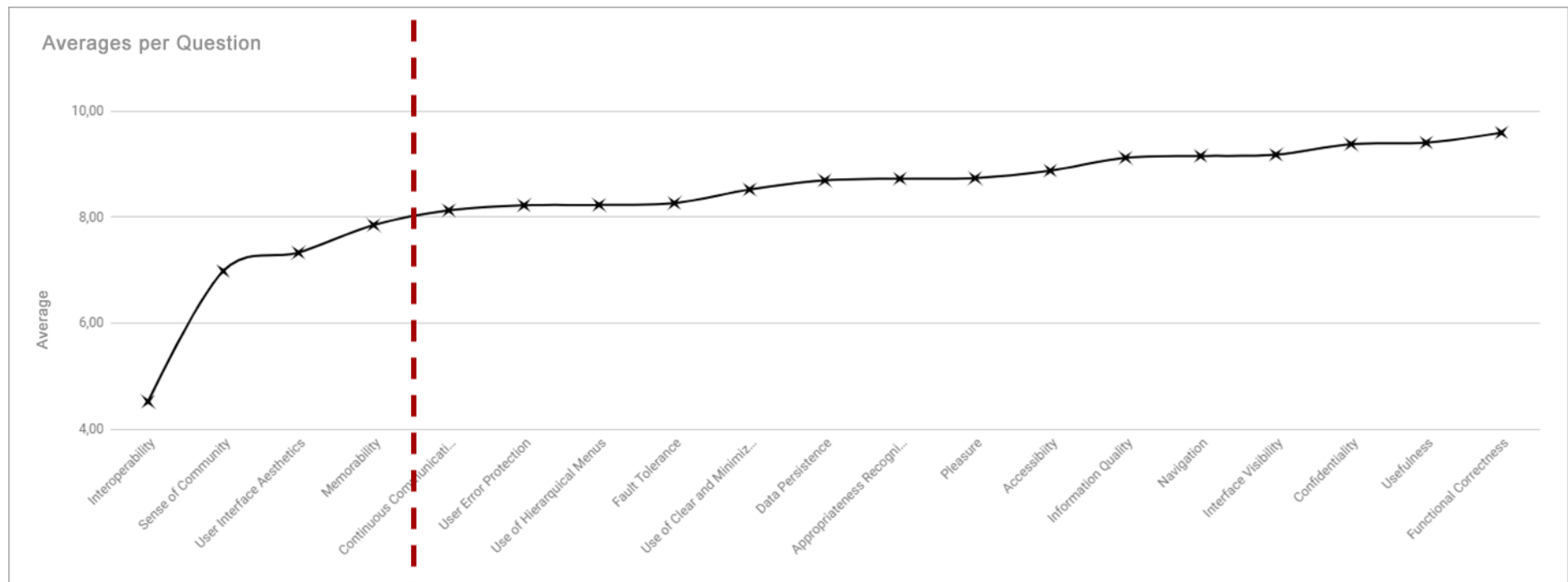


Figure 4.8 – Overall Average of the quality characteristics

**Confidentiality** and **Information Quality** are important for the **banking** app category due to the nature of the functionalities, which must be safe and accurate. **Interface Visibility** is also important because the user must easily find the required operations. **User Interface Aesthetics** also had a low evaluation for this app category, probably for the same reason of the **messaging** apps: the functionality matters more than the appearance. The **banking** category was also the one in which **Memorability** was considered the most essential, probably because these apps are opened at least once a month.

The overall averages for each quality sub-characteristics considering all 500 answers altogether are presented in Figure 4.8. The least essential characteristics were **Interoperability**, **Sense of Community**, **Memorability** and **User Interface Aesthetics**. The most essential characteristics were **Functional Correctness**, **Usefulness**, **Confidentiality**, **Interface Visibility**, **Navigation** and **Information Quality**.

#### 4.6 Threats to Validity

Threats should be dealt with care not to compromise the validity of the results. The classification presented by (Wohlin et al., 2012) will be followed, which considers internal validity, external validity, construct validity and conclusion validity.

Conclusion validity is concerned with the relationship between the treatment and the outcome. The 500 respondents permitted the inferring of some results, but none of them can be proven. The survey also does not contain hypothesis and was developed based on the goal of the study itself: the essentiality of quality characteristics. In addition, the results might be different if the survey was executed in another country. The decision of using 8.0 as the cutoff score based on the average of every answer regardless of the app category might also have biased the conclusion.

Construct validity is concerned with the relation between theory and observation. The VAS questions may not have been equivalent to the corresponding definition of the quality characteristic. Some results indicate the existence of bias due to the respondents not understanding the questions or understanding something different. For instance, **Pleasure's** low average for the **tourism** app category. In addition, **Use of Clear and Minimized Forms'** high average for the **messaging** app category and low average for the **tourism** app category. These results might indicate a misinterpretation of the respondents or some factor which biased the results. A certain quality characteristic may have been

considered not essential, when actually the respondents have understood something completely different.

Internal validity indicates that the treatment causes the outcome. The VAS questions were developed as simple as possible, yet some of them might have been subjective. In addition, the survey do not inquire the respondents about the apps which were first thought when the app category was selected. This extra question might have helped the interpretation of the results. Although, given the obtained data, most of the results make sense, according to the size of the sample.

External validity is concerned with generalization. Apart from the possibility of bias generated by the ambiguity of questions, some respondent might also have misunderstood the meaning of what was being inquired. In such case, the answer might have been different if the respondents indeed had understood the question.

#### **4.7 Final List of Essential Sub-characteristics**

The survey's motivation was to support the results of the systematic mapping and to confirm the essentiality of the quality sub-characteristics and attributes in Table 4.5 and Table 4.6. The resulting average of every question for every app category was 8.3. Given this calculation, it was decided to use the rounded value 8.0 as a cutoff score, in order to define what would be essential or not given the results of the survey. The crimson line in Figure 4.8 represents the cutoff score.

By applying this decision, **Interoperability**, **Sense of Community**, **Memorability** and **User Interface Aesthetics** were dismissed and not considered essential. The attributes **Information Quality**, **Navigation**, **Interface Visibility**, **Use of Clear and Minimized Forms**, **Use of Hierarchical Menus**, **Data Persistence** and **Continuous Communication** were considered essential. It is also important to remember that the quality sub-characteristics in Table 4.3 were considered essential before the conduction of the survey.

Besides dismissing the non-essential quality sub-characteristics, it is necessary to attach these attributes to the known quality characteristics in order to obtain a customized version of ISO/IEC 25010 quality models.

The attributes were attached to the quality models. **Navigation**, **Interface Visibility**, **Use of Clear and Minimized Forms** and **Use of Hierarchical Menus** are

directly related to **Usability**, so they were attached to **Usability** as sub-characteristics. **Data Persistence** and **Information Quality** are strongly related to the application's **Reliability**, so they were attached to **Reliability** as sub-characteristics. **Continuous Communication** is a quality in use attribute. None of the existent quality in use characteristics seemed similar enough to **Continuous Communication**, so it was considered a quality characteristic itself. Table 4.7 presents the final list of essential characteristics and sub-characteristics for the context of mobile applications, considering the customized arrangements. Figure 4.9, which is based on the methodology proposed in (Franch & Carvallo, 2003), presents customized versions of ISO/IEC 25010 quality models, considering both the dismissed and the newly added characteristics. In order to simplify the figure's visualization, the characteristics were represented by an identifying id, presented in Table 4.7 to Table 4.9.

Table 4.7 – Final list of essential characteristics for the context of mobile applications

Quality Model	Quality Characteristic	Quality Sub-characteristic	Reason for the essentiality Occurrences in the Mapping	Survey Overall Average
Product Quality	Functional Suitability (F)	Functional Correctness (F1)		9.6
	Performance Efficiency (P)	Time Behavior (P1)	9	
		Resources Utilization (P2)	10	
	Usability (U)	Learnability (U1)	11	
		Operability (U2)	15	
		Appropriateness Recognisability (U3)		8.7
		User Error Protection (U4)		8.2
		Accessibility (U5)		8.9
		Navigation (U6)		9.1
		Interface Visibility (U7)		9.2
		Use of Clear and Minimized Forms (U8)		8.5
		Use of Hierarchical Menus (U9)		8.2
	Reliability (R)	Fault Tolerance (R1)		8.3
		Information Quality (R2)		9.1
		Data Persistence (R3)		8.7
	Security (S)	Confidentiality (S1)		9.4
	Portability (Po)	Adaptability (Po1)	8	
Quality in Use	Satisfaction (Sa)	Usefulness (Sa1)		9.4
		Pleasure (Sa2)		8.7
	Effectiveness (E)	Effectiveness (E1)	9	
	Efficiency (Ef)	Efficiency (Ef1)	13	
	Context Coverage (C)	Context Completeness (C1)	9	
		Flexibility (C2)	10	



Quality Model	Quality Characteristic	Quality Sub-characteristic	Reason for the essentiality Occurrences in the Mapping	Survey Overall Average
	Continuous Communication (Co)	-		8.1

Table 4.8 – Dismissed quality characteristics and their ids in Figure 4.9

Quality Characteristic	ID
Compatibility	<b>Cm</b>
Freedom from Risk	<b>Fr</b>
Maintainability	<b>Ma</b>

Table 4.9 – Dismissed quality sub-characteristics, and their ids in Figure 4.9

Quality Sub-Characteristic	ID
Economic Risk Mitigation	<b>Fr1</b>
Environmental Risk Mitigation	<b>Fr2</b>
Health Risk Mitigation	<b>Fr3</b>
Functional Appropriateness	<b>F2</b>
Functional Completeness	<b>F3</b>
Comfort	<b>Sa3</b>
Trust	<b>Sa4</b>
Capacity	<b>P3</b>
Instalability	<b>Po2</b>
Replaceability	<b>Po3</b>
Coexistence	<b>Cm1</b>
Interoperability	<b>Cm2</b>
Maturity	<b>R4</b>
Availability	<b>R5</b>
Recoverability	<b>R6</b>
Modularity	<b>Ma1</b>
Reusability	<b>Ma2</b>
Analysability	<b>Ma3</b>
Modifiability	<b>Ma4</b>
Testability	<b>Ma5</b>
Integrity	<b>S2</b>
Non-Repudiation	<b>S3</b>
Accountability	<b>S4</b>
Authenticity	<b>S5</b>
User Interface Aesthetics	<b>U10</b>



Figure 4.9 – Customized versions of ISO/IEC 25010 quality models

## 4.8 Conclusion

This chapter presented the development and the execution of a survey intended to confirm the results of the previously conducted systematic mapping. An instrument was developed in order to obtain mobile users' opinion over a set of sub-characteristics. Primarily, a pilot test was conducted, which led to the instrument's adaptation.

The survey permitted the interpretation of results regarding both the quality sub-characteristics and characterization variables such as **age**, **gender**, **level of education** and favorite app category. The interpretation permitted the definition of criteria for deciding if a quality characteristic is essential or not.

A final list of essential quality characteristics was elaborated, merging results from the systematic mapping and the survey. The ISO/IEC 25010 quality model was also customized to reflect the results.

The next chapter will present the appraisal of a mobile application, based on the final list of essential quality characteristics and the QPS reference model.

# 5 Quality Evaluation of Mobile Applications

This chapter provides a detailed description of the dissertation proposal relating to the quality evaluation of mobile applications. An appraisal of a Brazilian mobile banking application was carried out to illustrate the suitability of the proposal. Section 5.1 briefly describes the QPS process and the appraisal method used for its verification. Section 5.2 shows an adaptation of the QPS method, which will therefore enable the appraisal of mobile applications. Section 5.3 describes the mobile banking application's appraisal and their outcome. Finally, section 5.4 presents the conclusion.

## 5.1 The QPS process and its appraisal method

QPS (Rocha et al., 2017) is a Brazilian reference model used for evaluating software products, described in Chapter 2. The QPS model assesses software products considering four dimensions: organizational dimension, software engineering dimension, services dimension and product quality dimension. Besides the appraisal of the product itself, the method also comprises the presentation of various documents issued by the company in charge of the product. These documents are referred to during the assessment of the Organizational, Software Engineering and Service dimensions.

The QPS appraisals method conforms to the ISO/IEC 33020 (ISO/IEC, 2015). It is a continuous assessment system which delivers results in a three-level ranking system: gold, silver and bronze. Progress is incremental: in order to reach a level, it is mandatory to comply with every requirement stated at the previous levels.

The QPS appraisal session starts with a diagnostic analysis of the product as a whole, whereby the product is assessed considering all the gold level requirements. Based on the diagnostic analysis results, the company in charge of the product determines the final appraisal level. In setting the gold level requirements as a starting point, a more wide-ranging report can be prepared, which will prove invaluable to the company and capable of guiding the continuous improvements of the product. The final appraisal session is held after a period of adjustments, considering the company's chosen level.

The evaluation team is composed of at least two certified evaluators, qualified to perform the QPS appraisals. One of them is assigned with the role of leader appraiser.

Members of the company in charge of the product are forbidden from joining the appraisal team, nevertheless.

To ensure the fairness of the QPS appraisals, the whole team shall have no connection with the organization or the product in question, providing adequate third-party services. Figure 5.1 presents the activities and tasks performed in the initial stage of the diagnostic appraisal process. Figure 5.2 presents the activities and tasks performed during the final stage of the appraisal process.

<p><b>Plan the Initial Diagnosis</b></p> <ul style="list-style-type: none"> <li>Report the appraisal to QPS's steering committee</li> <li>Start the appraisal planning</li> <li>Send the appraisal plan to the organizational unit in charge of the product</li> <li>Fill out the evaluation plan with organizational unit data and product data</li> <li>Complete the appraisal planning</li> </ul> <p><b>Gather data for the Initial Diagnosis</b></p> <ul style="list-style-type: none"> <li>Send the appraisal spreadsheet model to the organizational unit</li> <li>Fill out the appraisal spreadsheet</li> </ul> <p><b>Validate data from the Initial Diagnosis</b></p> <ul style="list-style-type: none"> <li>Conduct a kickoff meeting over initial diagnosis</li> <li>Demonstrate the product</li> <li>Present the processes</li> <li>Conduct the initial diagnosis</li> <li>Present the Initial Diagnosis's Report to the organizational unit</li> </ul>
--

Figure 5.1 –Initial Diagnosis Activities and Tasks

<p><b>Plan the Final Assessment</b></p> <ul style="list-style-type: none"> <li>Plan the Final Assessment</li> </ul> <p><b>Gather data for the Final Assessment Stage (if applicable)<sup>14</sup></b></p> <ul style="list-style-type: none"> <li>Make adjustments (if applicable)</li> </ul> <p><b>Validate data in the Final Assessment Stage</b></p> <ul style="list-style-type: none"> <li>Conduct a kickoff meeting to discuss the final assessment</li> <li>Present the performed adjustments</li> <li>Verify adjustments</li> <li>Conduct interviews (if applicable)</li> </ul> <p><b>Determine the Results</b></p> <ul style="list-style-type: none"> <li>Conduct initial characterization</li> <li>Confirm the initial characterization in a consensus meeting</li> <li>Characterize the level of achievement in the quality dimensions</li> <li>Assign level to the product</li> <li>Present preliminary results to the organizational unit</li> <li>Review characterization<sup>15</sup></li> </ul> <p><b>Report Results</b></p> <ul style="list-style-type: none"> <li>Appraisal result are reported to the sponsor</li> <li>Evaluation of the appraisal process is made by evaluation sponsor</li> <li>Evaluation of the appraisal process is made by the evaluation team</li> </ul>
--

<sup>14</sup> This task is performed where adjustments have proven to be necessary between the initial diagnosis stage and the final assessment stage.

<sup>15</sup> The presentation of preliminary results enables the organizational unit to provide feedback on the results to the evaluation team. In some cases, the supply of more information to the evaluators proves to be necessary, in order to enhance the team's understanding of the product. This information might change the characterizations.

Appraisal results are reported to the organizational unit Report and appraisal results are produced Audit appraisal documentation and publish results are stored
---

Figure 5.2 – Final Assessment Activities and Tasks

Each evaluator should conduct an individual initial characterization of all the four quality dimensions, thus attributing a level (gold, silver or bronze) in order to determine the results. Figure 5.3 describes the procedure for the initial characterization in case of the need for appraisal of existing products already on the market.

**Procedure for the initial characterization in case of appraisal of products already on the market**

In the **organizational dimension**, the evaluator must:

- i. Characterize the product documentation by assigning T (totally compliant), L (largely compliant), P (partially compliant) or N (noncompliant).
- ii. Characterize the degree of implementation of each process based on its expected results by assigning T (totally compliant), L (largely compliant), P (partially compliant) or N (noncompliant) to each instance and by defining the aggregate characterization for the product according to Table 5.2.
- iii. Characterize the degree of implementation of the processes' attributes by assigning T (totally compliant), L (largely compliant), P (partially compliant) or N (noncompliant) to each process attribute.

In the **software engineering dimension**, the evaluator must:

- i. Characterize the degree of implementation of each process according to its expected result by assigning T (totally compliant), L (largely compliant), P (partially compliant) or N (noncompliant) to each instance and by defining the aggregate characterization results for the product according to Table 5.2.
- ii. Characterize the degree of implementation of the processes' attributes by assigning T (totally compliant), L (largely compliant), P (partially compliant) or N (noncompliant) to each process attribute.

In the **services dimension**, the evaluator should:

- i. Characterize the degree of implementation of each process according to its expected results by assigning T (totally compliant), L (largely compliant), P (partially compliant) or N (noncompliant) to each instance and by defining the aggregate characterization results for the product according to Table 5.2.
- ii. Characterize the degree of implementation of the processes' attributes by assigning T (totally compliant), L (largely compliant), P (partially compliant) or N (not compliant) to each process attribute.

In the **product quality dimension**, the evaluator must:

- i. Characterize the product's description by assigning T (totally compliant), L (largely compliant), P (partially compliant) or N (noncompliant).
- ii. Characterize the user documentation by assigning T (totally compliant), L (largely compliant), P (partially compliant) or N (notncompliant).
- iii. Characterize the degree of compliance of the quality measures by assigning T (totally compliant), L (largely compliant), P (partially compliant) or N (noncompliant).

**In order to assign T, L, P or N, the evaluator should follow the rules described in Table 5.1.**

**In order to characterize the product's result, the evaluator should follow the aggregation rules in Table 5.2.**

Figure 5.3 – Procedure for the initial characterization of products already on the market

The characterization rules are defined in Table 5.1 and the aggregation rules for instance target characterization (projects or services) and characterization of product requirement are defined in Table 5.2.

Table 5.1 – Characterization Rules

Level of compliance		Description	Percentage of compliance
T	Totally compliant	There exists sufficient and suitable evidence to demonstrate the complete fulfillment of the reference model's requirement for the product.	>85% a 100%
L	Largely compliant	There exists sufficient and suitable evidence to demonstrate a significant degree of commitment to the reference model's requirement for the product. There exists one or more weak points related to this requirement in the reference model, but it does not compromise the compliance with the requirement.	>50% a 85%
P	Partially compliant	There exists little suitable evidence to demonstrate partial fulfillment of the reference model's requirement for the product. There exists one or more weak points related to this requirement in the reference model which do compromise the compliance with the requirement.	>15% a 50%
N	Non Compliant	There exists little or no evidence to demonstrate partial commitment to the reference model's requirement for the product.	0 a 15%

Table 5.2 – Aggregation rules

Characterization of instances (projects or services)	Product Characterization	Notes
all X (i.e. all T or all L or all P or all N)	X	If the requirement characterization is the same for each project or service, this will also be the product characterization.  If the requirement characterization is NA given the stage of development of a product or service, the characterization remains unaffected.
all T or all L	L	If the requirement characterization is T or L for each project or service, the product characterization will be L.  If the requirement characterization is NA given the stage of development of a product or service, the characterization remain unaffected.
P exists, but not N	L or P	The evaluation team decides by consensus.
N exists	N, P or L	The evaluation team decides by consensus.

A quality dimension will be rated as bronze level if all the dimension requirements for the bronze level were satisfactorily met. A quality dimension will be rated as silver

level if all the dimension requirements for the bronze and silver levels were satisfactorily met. A quality dimension will be rated as gold level if all the dimension requirements for the bronze, silver and gold levels were satisfactorily met.

After the completion of the initial characterization, the results are confirmed in a consensus meeting, involving the whole appraisal team.

Finally, the process is completed with the product characterization. The product will be rated as bronze level if all quality dimensions achieved at least bronze-level status. The product will be rated as silver level if all quality dimensions achieved at least silver-level status. The product will be characterized with gold level if all quality dimensions achieved at least gold-level status.

QPS appraised four products so far:

- WTS Corporate: travel management software designed by the company Monteiro e Gutierrez Sistemas Ltda, based in Rio de Janeiro.
- Pirâmide: ERP designed by the company PROCENGE, based in Recife.
- Estoque SQL: software designed by the company Nasajon, based in Rio de Janeiro.
- RSI: a management tool for institutional services designed by Fiocruz, based in Rio de Janeiro.

## **5.2 Appraisal of Mobile Applications**

This dissertation aims to evaluate the quality of mobile applications and propose an effective means of assessing them. To achieve this objective, a systematic mapping was conducted (described in Chapter 3) as well as a survey with mobile users to verify the results of the mapping (described in Chapter 4). This section describes a proposition for appraising mobile applications, based on QPS reference model and its appraisal process and method. Considering that QPS already evaluates products by means of quality characteristics in the Product Quality Dimension, and also considering that at gold level the model evaluates context-specific quality characteristics, it was decided to use QPS as a basis for assessing the quality of mobile applications. It is consonant with the needs of this work.

Mobile applications may be appraised in two cases:

1) **The appraisal is commissioned by the company in charge of the product.**

The company wishes to conduct a thorough appraisal, considering all QPS quality dimensions. In this setting, the appraisal process is completely executed, following the Initial Diagnosis and the Final Assessment procedures.

2) **The appraisal is based on a mobile application available in an app store,**

where no documentation is available. In this setting, only the product quality dimension can be appraised. In addition, provided that the company in charge of the product does not intervene with the appraisal process and consequently does not perform adjustments, only the Final Assessment is conducted. Figure 5.4 presents the activities and tasks in this case.

<b>Plan the Final Appraisal</b>
Plan the Final Appraisal
<b>Validate data in the Final Appraisal</b>
Conduct a kickoff meeting to discuss the final appraisal
Appraise the product
<b>Determine the Results</b>
Conduct initial characterization
Confirm the initial characterization in a consensus meeting
Assign level to the product quality dimension
<b>Report Results</b>
Evaluate the evaluation team's appraisal execution
Generate appraisal's final report
Audit appraisal
Store documentation and publish result

Figure 5.4 – Activities and Tasks used in the Final Appraisal (apps in an app store)

In the first case, in which the complete QPS appraisal is executed, the category of software product does not affect the organizational dimension, the software engineering dimension or the services dimension, so these dimensions will not be described here. This thesis intends to define a procedure for appraising the product quality dimension only.

A list of quality sub-characteristics for appraising mobile applications in the product quality dimension was drawn up. This list contains the quality sub-characteristics already present in QPS for any software product and the sub-characteristics assembled in Table 4.7.

Interoperability, which was originally considered not essential given the results of the survey, was also added to the list. The survey was presented at the Brazilian Symposium on Software Quality (SBQS 2019) and the participants clearly showed that



removing Interoperability from the mobile application appraisal process might jeopardize the results.

Table 5.3 lists questions regarding the product quality sub-characteristics, which ought to be considered in the appraisal phase. It should be noticed that Maintainability sub-characteristics cannot be appraised in the second case and therefore must be tagged as “not evaluable”. Table 5.4 lists the quality in use sub-characteristics to be considered in the appraisal phase. The questions in these tables are identified by means of the following rules:

- Alphabetical abbreviation: which represents the related quality characteristic.
- Sequential number related to the quality characteristic.
- G (Generic) means that the question should always be considered in the appraisal and S (Specific) means that the question is only referent to mobile applications.

Table 5.3 – Questions for the assessment of product quality sub-characteristics

<b>Id</b>	<b>Quality Sub-characteristic</b>	<b>Question</b>
US1-G	Operability	Are the task behavior and appearance consistent?
US2-G	Operability	Are the product's messages clear?
US3-G	Operability	Are there features such as undo, or at least a confirmation dialog for fundamental tasks?
US4-G	User Interface Aesthetics	Is the appearance of the user interface pleasant?
US5-S	Appropriateness Recognizability	When the mobile application is being used for the first time, are the users able to realize if it is adequate for their expected needs?
US6-S	Learnability	Is it easy to learn how to use the mobile application functionalities?
US7-S	Operability	Is it easy to operate and control the mobile application in order to use its functionalities?
US8-S	User Error Protection	Does the mobile application prevent the user from committing mistakes?
US9-S	Accessibility	Does the mobile application contain adjustments that allow the participation of users with motor, visual and auditory impairments?
US10-S	Navigation	Is it easy to find the functionality or information that the user needs?
US11-S	Interface Visibility	Are the screens organized in such a way that facilitates a fast understanding of the texts?
US12-S	Use of Clear and Minimized Forms	Does the mobile application contain clear forms with help?
US13-S	Use of Hierarchical Menus	Does the mobile application contain enough menus, and are they simple and easy to navigate?

<b>Id</b>	<b>Quality Sub-characteristic</b>	<b>Question</b>
SE1-G	Confidentiality	Does the product contain access control against unauthorized access?
SE2-S	Confidentiality	Does the mobile application request the user's permissions before using information?
SE3-G	Integrity	Are there preventive measures to avoid that unauthorized access corrupt or modify user data?
AF1-S	Functional Correction	Does the mobile application do exactly what the user expects?
ED1-S	Time Behavior	Is the mobile application response time adequate?
ED2-S	Resources Utilization	Is the mobile application use of memory and battery resources adequate?
COMP1-S	Interoperability	Does the mobile application satisfactorily communicate data among other applications?
CONF1-S	Fault Tolerance	Does the mobile application properly behave even in case of software or device problems?
CONF2-S	Data Persistence	After possible system pauses, does the mobile application store appropriate information?
CONF3-S	Information Quality	Does the mobile application provide correct and adequate information to satisfy the user's needs?
PORT1-S	Adaptability	Is the mobile application adapted to different platforms?
MAN1-G	Analyzability	Does the traceability structure exist, and is it complete to support changes' impact analysis?
MAN2-G	Testability	Does the traceability structure exist, and is it complete to support the run of tests after changes?
MAN3-G	Testability	Are test cases available for conducting regression tests after changes are made?

Table 5.4 – Questions for the evaluation of quality in use sub-characteristics

<b>Id</b>	<b>Quality Sub-characteristic</b>	<b>Question</b>
EFI1-S	Efficiency	When using the mobile application, does the user achieve their objectives without getting tired?
EFE1-S	Effectiveness	When using the mobile application, does the user achieve their objectives with correctness and completeness?
SAT1-S	Usefulness	Is the mobile application useful for the users to achieve the objectives which made them use it?
SAT2-S	Pleasure	Is the mobile application pleasant to use?
COB1-S	Context Coverage	Does the mobile application work properly in every expected context?
COB2-S	Flexibility	Does the mobile application work properly in contexts of use other than those expected?
COM1-S	Continuous Communication	Is the mobile application usage not disturbed by problems like poor internet connection or weak GPS signal?

The evaluation procedure consists of three steps: characterizing quality sub-characteristics, characterizing quality characteristics and attributing a level.

- **Characterization of quality sub-characteristics:** Primarily, the evaluators should conduct an individual assessment of each sub-characteristic using the questions of the appraisal instrument. Then the evaluators should reach a consensus concerning each question related to the sub-characteristic and rate each of them T (totally compliant), L (largely compliant), P (partially compliant) or N (noncompliant). Some might also be considered “not evaluable”.
- **Characterization of Quality Characteristics:** After the characterization of each question related to sub-characteristics, the ratings T, L, P and N should be mapped to values 3, 2, 1 or 0, respectively. “Not evaluable” sub-characteristics should be ignored. The values related to the same characteristics should be grouped and their median should be calculated<sup>16</sup>. The quality characteristics will be given a rating of T, L, P or N depending on the following rules<sup>17</sup>:
  - T if median ranges from 2.55 to 3
  - L if median ranges from 1.5 to 3
  - P if median ranges from 0.45 to 1.5
  - N if median less than 0.45
- **Level Assignment:** The product quality dimension will be awarded a level (gold, silver or bronze) depending on the ratings obtained by characteristics and sub-characteristics, following these rules:
  - Bronze level:
    - The quality sub-characteristic with ID SE1-G was characterized with T and the other quality sub-characteristics were given a T or L.
  - Silver level:
    - The quality sub-characteristic with id SE1-G was given a T and the other quality characteristics were given a T or L.
    - Every quality in use characteristic was given either T or L.
  - Gold level:
    - Every generic quality sub-characteristic (whose ID ends in G) was given a T.

<sup>16</sup> The decision of using median to characterize quality characteristics is based on the study in (Idri et al., 2017).

<sup>17</sup> This characterization is in accordance with the rules in Table 5.1.

- Every quality in use characteristic was given either T or L.
- Every product quality characteristic was given either T or L.

The appraisal team shall use an appraisal instrument document (Annex B) containing instructions for characterization and level assignment and the questions presented in Table 5.3 and Table 5.4.

### **5.3 Appraisal Planning and Execution for a Mobile Applications**

An appraisal was conducted with a Brazilian mobile banking application, in order to illustrate the feasibility of the thesis proposal. This mobile application is available in an app store and the company was not involved in the process. Consequently, this scenario describes the second appraisal case, described in the beginning of section 5.2. Moreover, in this case only the Final Assessment is pertinent. Three QPS documents were adapted for the appraisal: The Appraisal Plan, Product Evaluation Final Report and Evaluation of the Evaluation Process by the Team of Evaluators.

#### **5.3.1 Appraisal Plan**

The iOS version of a Brazilian mobile banking application was selected for this appraisal. The appraisal execution was scheduled to take place in one morning. Two certified QPS evaluators were selected to make up the evaluation team and one of them was assigned as the leader appraiser. The appraisal planning was conducted by the local coordinator (the author of this thesis) and by the leader appraiser, as requested by QPS. The Appraisal Plan is in Annex C.

#### **5.3.2 Execution of the Appraisal**

As defined in the schedule presented in the Appraisal Plan, the first activity was the kick-off meeting. In this meeting, the local coordinator explained the evaluation instructions and the questions to the appraisal team. The team's doubts were solved by the local coordinator and as soon as everyone stated to have understood the questions and procedure, the team was left alone, and the appraisal started.

The appraisal lasted two hours. By the end, the lead appraiser produced the final report and both members of the appraisal team filled a document concerning the process execution, the method and the appraisal instrument. The tasks “Audit appraisal” and “Documentation and publish results are stored,” specified in Table 5.4 were not carried out since no publication would be prepared for this specific appraisal.

Table 5.5 and Table 5.6 display the ratings assigned to each question related to the product quality sub-characteristics and to the quality in use sub-characteristics, respectively.

Table 5.5 – Characterization of the questions of the product quality sub-characteristics

<b>Id</b>	<b>Quality Sub-characteristic</b>	<b>Question</b>	<b>Characterization</b>
US1-G	Operability	Are the task behavior and appearance consistent?	<b>T</b>
US2-G	Operability	Are the product's messages clear?	<b>T</b>
US3-G	Operability	Are there features such as undo, or at least a confirmation dialog for fundamental tasks?	<b>T</b>
US4-G	User Interface Aesthetics	Is the appearance of the user interface pleasant?	<b>T</b>
US5-S	Appropriateness Recognizability	When the mobile application is being used for the first time, are the users able to realize if it is adequate for their expected needs?	<b>L</b>
US6-S	Learnability	Is it easy to learn how to use the mobile application functionalities?	<b>T</b>
US7-S	Operability	Is it easy to operate and control the mobile application in order to use its functionalities?	<b>T</b>
US8-S	User Error Protection	Does the mobile application prevent the user from committing mistakes?	<b>L</b>
US9-S	Accessibility	Does the mobile application contain adjustments that allow the participation of users with motor, visual and auditory impairments?	<b>N</b>
US10-S	Navigation	Is it easy to find the functionality or information that the user needs?	<b>L</b>
US11-S	Interface Visibility	Are the screens organized in such a way that facilitates a fast understanding of the texts?	<b>T</b>
US12-S	Use of Clear and Minimized Forms	Does the mobile application contain clear forms with help?	<b>T</b>
US13-S	Use of Hierarchical Menus	Does the mobile application contain enough menus, and are they simple and easy to navigate?	<b>L</b>
SE1-G	Confidentiality	Does the product contain access control against unauthorized access?	<b>T</b>
SE2-S	Confidentiality	Does the mobile application request the user's permissions before using information?	<b>T</b>
SE3-G	Integrity	Are there preventive measures to avoid that unauthorized access corrupt or modify user data?	<b>NA</b>
AF1-S	Functional Correction	Does the mobile application do exactly what the user expects?	<b>T</b>
ED1-S	Time Behavior	Is the mobile application response time adequate?	<b>L</b>
ED2-S	Resources Utilization	Is the mobile application use of memory and battery resources adequate?	<b>L</b>
COMP1-S	Interoperability	Does the mobile application satisfactorily communicate data among other applications?	<b>L</b>
CONF1-S	Fault Tolerance	Does the mobile application properly behave even in case of software or device problems?	<b>L</b>
CONF2-S	Data Persistence	After possible system pauses, does the mobile application store appropriate information?	<b>T</b>
CONF3-S	Information Quality	Does the mobile application provide correct and adequate information to satisfy the user's needs?	<b>T</b>

<b>Id</b>	<b>Quality Sub-characteristic</b>	<b>Question</b>	<b>Characterization</b>
PORT1-S	Adaptability	Is the mobile application adapted to different platforms?	<b>NA</b>
MAN1-G	Analyzability	Does the traceability structure exist, and is it complete to support changes' impact analysis?	<b>NA</b>
MAN2-G	Testability	Does the traceability structure exist, and is it complete to support the run of tests after changes?	<b>NA</b>
MAN3-G	Testability	Are test cases available for conducting regression tests after changes are made?	<b>NA</b>

Table 5.6 – Characterization of the questions of quality in use sub-characteristics

<b>Id</b>	<b>Quality Sub-characteristic</b>	<b>Question</b>	<b>Characterization</b>
EFI1-S	Efficiency	When using the mobile application, does the user achieve their objectives without getting tired?	<b>T</b>
EFE1-S	Effectiveness	When using the mobile application, does the user achieve their objectives with correctness and completeness?	<b>T</b>
SAT1-S	Usefulness	Is the mobile application useful for the users to achieve the objectives which made them use it?	<b>T</b>
SAT2-S	Pleasure	Is the mobile application pleasant to use?	<b>T</b>
COB1-S	Context Completeness	Does the mobile application work properly in every expected context?	<b>T</b>
COB2-S	Flexibility	Does the mobile application work properly in contexts of use other than those expected?	<b>NA</b>
COM1-S	Continuous Communication	Is the mobile application usage not disturbed by problems like poor internet connection or weak GPS signal?	<b>N</b>

Table 5.7 and Table 5.8 display the resulting ratings of the product quality characteristics and of the quality in use characteristics, respectively. Portability and Maintainability could not be appraised.

Table 5.7 – Characterization of the product quality characteristics

<b>Quality Characteristic</b>	<b>Characterization</b>
Usability	<b>T</b>
Security	<b>T</b>
Functional Suitability	<b>T</b>
Performance Efficiency	<b>L</b>
Compatibility	<b>L</b>
Reliability	<b>T</b>
Maintainability	<b>NA</b>
Portability	<b>NA</b>

Table 5.8 – Characterization of the product quality characteristics

Quality Characteristic	Characterization
Efficiency	T
Effectiveness	T
Satisfaction	T
Context Coverage	T
Continuous Communication	N

The level assignment rules described in section 5.2 were then applied, the results being shown in Table 5.5 to Table 5.8. The iOS version of the banking application was awarded the BRONZE level. Notice that except for the characterization of question COM1 the product would have been awarded the silver level. Even though question US9-S did not affect the result, it was also rated with N, indicating no concern about accessibility. Annex D contains the Appraisal Final Report filled by the leader appraiser.

### 5.3.3 Appraisal Team's Feedback concerning the Appraisal

After concluding the appraisal, the appraisal team filled an individual document concerning the appraisal process execution, the method and the appraisal instrument. This document inquired about improvement suggestions for both the instrument and the appraisal method. The following questions were raised:

1. Due to the similarities between the words characteristics and characterization, the titles became rather confusing.
2. Even though median is a simple measure, it might not be the best measure for this type of evaluation. One evaluator suggested weighted arithmetic mean.
3. COB1, COB2 and CONF1-S demands extra examples due to their complexity.
4. One evaluator suggested a joint assessment since the beginning of the process, in order to facilitate the consensus meeting.
5. The mobile application details (name, version and description) are specified in the Final Report but should also be included in the Appraisal Instrument.
6. The text areas for weak points, strong points and improvement suggestions are specified in the Final Report but should also be included in the Appraisal Instrument.

Annex E contains the revised instrument. The terms "characteristics" and "characterization" cannot be changed for synonyms due to the same use in QPS, so the

titles were not changed. The questions with IDs COB1, COB2 and CONF1-S were extended with examples. The mobile application details and the weak points, strong points and improvement suggestions text areas were added to the instrument.

The suggestion regarding the conduction of a joint assessment since the beginning of the appraisal would not affect the instrument, only the method itself. However, we would rather keep the process as similar as possible to QPS appraisal method.

The study in (Idri et al., 2017) describes an appraisal of four quality characteristics in pregnancy mobile monitoring applications, by conducting a questionnaire with Likert questions. In this study, the median was used to calculate the degree of each quality characteristic in each app. We chose to use median as well, due to the similarities with the assessment described in this chapter. We found no studies in literature sharing similarities and using weighted arithmetic means.

## **5.4 Conclusion**

This chapter presented an appraisal of a Brazilian mobile banking application. The objective of the appraisal was to illustrate and evaluate the use feasibility of the dissertation's proposition. The appraisal was based on the QPS, a Brazilian reference model for appraising software products. The QPS method and procedure were adapted with extra rules and an instrument was defined for the assessment of mobile applications.

The appraisal was conducted by a team of QPS certified evaluators and lasted two hours. By the end, the leader appraiser produced a final report of the appraisal and the whole team filled in an evaluation form of the appraisal instrument, process and method.

The appraisal was conducted as expected, without difficulties. The appraisal team suggested some improvements, which were applied to a newer version of the appraisal instrument. The evaluation satisfactorily showed the possibility of using the proposal of this dissertation as a basis to conduct a mobile application's appraisal.



## 6 Conclusion

This dissertation identified quality characteristics and sub-characteristics and proposes an appraisal procedure for mobile applications. Section 6.1 summarizes the obtained results. Section 6.2 presents the bibliographical production of the author during the development of the dissertation. Section 6.4 discusses future work which might derive from the results.

### 6.1 Obtained Results

Mobile applications are different from other software products, given its unique features and its mobile hardware. They shall be adapted to a variety of contexts by using accelerometer, GPS, Bluetooth and other sensors. They are always with the users, who possess a wide range of personal characteristics, so it is important to develop simple and easy to use applications. Every new generation of mobile devices integrates more possibilities of functionalities to their applications. Desktop and web systems are different, for instance, because they rely less on sensors and context variation, and have a much more stable internet. It is important to know in advance the stakeholders and the particularities of a software type.

The aim of the systematic mapping was to identify quality characteristics pertinent in the context of mobile applications, both those previewed in ISO/IEC 25010 and those not previewed. The amount of returned studies was considerably large, and the themes of each study was quite diverse. However, the systematic mapping allowed only the ranking of quality characteristics by occurrences. A survey was proven necessary in order to question mobile application users about their opinion over a set of sub-characteristics. With the result, it was possible to compose an adapted version of ISO/IEC 25010 quality models for mobile applications.

An evaluation procedure was defined, based on QPS reference model and on the adapted quality models. QPS defines criteria for appraising both general and specific quality characteristics of software products, so it was feasible to adapt it. The evaluation of the mobile banking application confirmed the feasibility of evaluating mobile applications by using the previously adapted quality models from ISO/IEC 25010. The evaluation was quite simple and the results were reasonable for what was intended,

indicating that the appraisal of mobile applications can be conducted by using the identified characteristics/sub-characteristics and the provided appraisal procedure.

## 6.2 Limitations

The boundaries of this study are as follows:

- The systematic mapping presented in Chapter 3 is limited to its threats to validity, listed in section 3.7. We highlight some of these threats:
  - Some papers provided quality characteristics with differing definitions from ISO/IEC 25010, which demanded a deep interpretation of what the authors meant.
  - The search string may not have captured other relevant papers.
  - The interpretation of the abstracts and the manual selection of papers via snowballing may have biased the results.
  - The research questions were subjective and demanded effort to both understand what to be extracted and how to interpret the extracted data.
  - The quantity of papers may have been relatively small, possibly due to a too restrictive search string.
- The survey presented in Chapter 4 is also limited to its threats to validity, listed in section 4.6. We highlight some of these threats:
  - The survey does not contain hypothesis and was developed based on the goal of the study itself.
  - The answers permitted the inferring of some results, but none of them can be proven due to the much greater size of the population.
  - The VAS questions were developed as simple as possible, yet some of them might have been subjective.
  - Some respondent might have misunderstood the meaning of what was being inquired.
- Some decisions were made by the author during the development of this work, and they might have biased the results:
  - Not including quality characteristics in the survey.
  - Only adding the quality sub-characteristics with occurrences between five and seven in the survey.
  - Using mean 8.0 as a cutoff score, given the results of the survey.

- Deciding where the new sub-characteristics should be appended in the adaption of the ISO/IEC 25010 quality models.
- Manually adding Interoperability sub-characteristic to the appraisal procedure, regardless of the survey results.

### **6.3 Bibliographic Production**

During the master's course, the author presented the progress of the research in two symposiums. One article was written and published.

- The dissertation proposal was presented in the Workshop on Software Quality Theses and Dissertations (WTDQS) during the Brazilian Software Quality Symposium 2018 (SBQS'18).
- The survey was presented in SBQS'19 and published in ACM library.
  - Vitor Maia, Taisa G. Gonçalves, Ana Regina Rocha. 2019. Quality Characteristics of Mobile Applications: A Survey in Brazilian Context. In Proceedings of the Brazilian Software Quality Symposium (SBQS'19). ACM, New York, NY, USA, 109-118. <https://doi.org/10.1145/3364641.3364654>

### **6.4 Future Work**

This research may be improved by means of future works. The improvement perspectives are listed below:

- The appraisal method shall be executed in different mobile applications.

The appraisal presented in Chapter 5 indicated the possibility of appraising mobile banking applications and potentiality other app categories sharing similar characteristics. However, it might also be interesting to appraise mobile applications with very different settings, e.g., applications which must be used in an open environment, or in movement.

- The appraisal instrument shall be further refined.

The appraisal presented in Chapter 5 was conducted with the assistance of an appraisal instrument. By conducting a greater number of appraisals, more improvement suggestions would be recommended, leading to a more refined appraisal instrument.

- The survey shall be rerun with different settings.

The survey was executed in Brazil and was closed when the 500th question was received. It might be interesting to rerun it with a broader audience, in order to check if the obtained results remain the same even when different cultures are involved. The survey might also be changed to comprise new questions: 1) Ask the respondents which app they thought about when they selected the app category, 2) Additional demographic questions.

- The appraisal method shall have five different settings.

The ISO/IEC 25010 quality models were adapted based on the list of essential quality characteristics for mobile applications. Although, as seen in Figure 4.6, the survey permitted five different results, one for each app category. Elaborating five specialized lists of essential quality characteristics, instead of only one, might affect positively the elaboration of the appraisal method.

- The criteria for selecting which quality sub-characteristics would compose the survey shall be based on quartiles.

For deciding which quality sub-characteristics should compose the survey, three intervals were defined based on the range of occurrences of sub-characteristics, which were comprised between zero and fifteen. Although, the limits of the intervals were manually selected, without statistical basis. In a future work, the occurrences of the sub-characteristics might be used as samples and the intervals might be defined based on quartiles. By removing the zeros and the characteristic with fifteen occurrences (Operability), which was an outlier, the first quartile shall be 1-3, the third quartile shall be 7-15 and the interval to be added to the survey shall be 4-6, instead of 5-7.

# Bibliographical References

- Abusair, M. (2017). User- and analysis-driven context aware software development in mobile computing. *Proceedings of the 2017 11th Joint Meeting on Foundations of Software Engineering - ESEC/FSE 2017*, 1022–1025. <https://doi.org/10.1145/3106237.3119873>
- Alaa, G., Menshawi, M., & Saeed, M. (2013). Design criteria and software metrics for efficient and effective web-enabled mobile applications. *2013 8th International Conference on Computer Engineering & Systems (ICCES)*, 295–300. <https://doi.org/10.1109/ICCES.2013.6707222>
- Alves, A., Salviano, C., Stefanuto, G., Maintinguer, S., Mattos, C., Zeitoum, C., Martinez, M., & Reuss, G. (2014). *CERTICS Assessment Methodology for Software Technological Development and Innovation*. 174–177. <https://doi.org/10.1109/QUATIC.2014.32>
- Aranha, E., & Borba, P. (2007). An Estimation Model for Test Execution Effort. *First International Symposium on Empirical Software Engineering and Measurement (ESEM 2007)*, 107–116. <https://doi.org/10.1109/ESEM.2007.73>
- Bachiri, M., Idri, A., Fernandez-Aleman, J. L., & Toval, A. (2015). A preliminary study on the evaluation of software product quality of pregnancy monitoring mPHRs. *2015 Third World Conference on Complex Systems (WCCS)*, 1–6. <https://doi.org/10.1109/ICoCS.2015.7483224>
- Balagtas-Fernandez, F., & Hussmann, H. (2009). A Methodology and Framework to Simplify Usability Analysis of Mobile Applications. *2009 IEEE/ACM International Conference on Automated Software Engineering*, 520–524. <https://doi.org/10.1109/ASE.2009.12>
- Baloh, M., Zupanc, K., Kosir, D., Bosnic, Z., & Scepanovic, S. (2015). A quality evaluation framework for mobile learning applications. *2015 4th Mediterranean Conference on Embedded Computing (MECO)*, 280–283. <https://doi.org/10.1109/MECO.2015.7181923>
- Barnett, S., Vasa, R., & Tang, A. (2015). A Conceptual Model for Architecting Mobile Applications. *2015 12th Working IEEE/IFIP Conference on Software Architecture*, 105–114. <https://doi.org/10.1109/WICSA.2015.28>
- Barney, S., & Wohlin, C. (2009). Software Product Quality: Ensuring a Common Goal. In Q. Wang, V. Garousi, R. Madachy, & D. Pfahl (Eds.), *Trustworthy Software*

- Development Processes* (Vol. 5543, pp. 256–267). Springer Berlin Heidelberg.  
[https://doi.org/10.1007/978-3-642-01680-6\\_24](https://doi.org/10.1007/978-3-642-01680-6_24)
- Basili, V. R., Caldiera, G., & Rombach, H. D. (1994). The Goal Question Metric Approach. In *Encyclopedia of Software Engineering*. Wiley.
- Ben Ayed, E., Kolski, C., Magdich, R., & Ezzedine, H. (2016). Towards a context based Evaluation Support System for Quality in Use assessment of mobile systems. *2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 004350–004355. <https://doi.org/10.1109/SMC.2016.7844915>
- Bezerra, C. I. M., Andrade, R. M. C., & Monteiro, J. M. S. (2014). Measures for Quality Evaluation of Feature Models. In I. Schaefer & I. Stamelos (Eds.), *Software Reuse for Dynamic Systems in the Cloud and Beyond* (Vol. 8919, pp. 282–297). Springer International Publishing. [https://doi.org/10.1007/978-3-319-14130-5\\_20](https://doi.org/10.1007/978-3-319-14130-5_20)
- Biolchini, J., Mian, P. G., Natali, A. C., & Travassos, G. H. (2005). *Systematic Review in Software Engineering: Relevance and Utility* (p. 31). PESC/COPPE/UFRJ. <http://www.cos.ufrj.br/uploadfiles/es67905.pdf>
- Boehm, B. W. (1978). *Characteristics of software quality*. North-Holland Pub. Co. <https://books.google.com.br/books?id=Cdm0AAAAIAAJ>
- Canfora, G., Di Sorbo, A., Mercaldo, F., & Visaggio, C. A. (2016). Exploring Mobile User Experience Through Code Quality Metrics. In P. Abrahamsson, A. Jedlitschka, A. Nguyen Duc, M. Felderer, S. Amasaki, & T. Mikkonen (Eds.), *Product-Focused Software Process Improvement* (Vol. 10027, pp. 705–712). Springer International Publishing. [https://doi.org/10.1007/978-3-319-49094-6\\_59](https://doi.org/10.1007/978-3-319-49094-6_59)
- Catolino, G. (2018). Does source code quality reflect the ratings of apps? *Proceedings of the 5th International Conference on Mobile Software Engineering and Systems - MOBILESoft '18*, 43–44. <https://doi.org/10.1145/3197231.3198447>
- Corral, L. (2012). Using software quality standards to assure the quality of the mobile software product. *Proceedings of the 3rd Annual Conference on Systems, Programming, and Applications: Software for Humanity - SPLASH '12*, 37. <https://doi.org/10.1145/2384716.2384734>
- Corral, L., & Fronza, I. (2015). Better Code for Better Apps: A Study on Source Code Quality and Market Success of Android Applications. *2015 2nd ACM International Conference on Mobile Software Engineering and Systems*, 22–32. <https://doi.org/10.1109/MobileSoft.2015.10>

- Corral, L., Sillitti, A., & Succi, G. (2014). Defining Relevant Software Quality Characteristics from Publishing Policies of Mobile App Stores. In I. Awan, M. Younas, X. Franch, & C. Quer (Eds.), *Mobile Web Information Systems* (Vol. 8640, pp. 205–217). Springer International Publishing. [https://doi.org/10.1007/978-3-319-10359-4\\_17](https://doi.org/10.1007/978-3-319-10359-4_17)
- Cortellessa, V., Mirandola, R., & Potena, P. (2010). Selecting Optimal Maintenance Plans Based on Cost/Reliability Tradeoffs for Software Subject to Structural and Behavioral Changes. *2010 14th European Conference on Software Maintenance and Reengineering*, 21–30. <https://doi.org/10.1109/CSMR.2010.15>
- Cortimiglia, M. N., Ghezzi, A., & Renga, F. (2011). Mobile Applications and Their Delivery Platforms. *IT Professional*, 13(5), 51–56. <https://doi.org/10.1109/MITP.2011.84>
- Dantas, V. L. L., Marinho, F. G., da Costa, A. L., & Andrade, R. M. C. (2009). Testing requirements for mobile applications. *2009 24th International Symposium on Computer and Information Sciences*, 555–560. <https://doi.org/10.1109/ISCIS.2009.5291880>
- de Sá, M., & Carriço, L. (2008). Lessons from early stages design of mobile applications. *Proceedings of the 10th International Conference on Human Computer Interaction with Mobile Devices and Services - MobileHCI '08*, 127. <https://doi.org/10.1145/1409240.1409255>
- de Souza, L. S., & de Aquino, G. S. (2014). Mobile Application Development: How to Estimate the Effort? In B. Murgante, S. Misra, A. M. A. C. Rocha, C. Torre, J. G. Rocha, M. I. Falcão, D. Taniar, B. O. Apduhan, & O. Gervasi (Eds.), *Computational Science and Its Applications – ICCSA 2014* (Vol. 8583, pp. 63–72). Springer International Publishing. [https://doi.org/10.1007/978-3-319-09156-3\\_5](https://doi.org/10.1007/978-3-319-09156-3_5)
- de Souza, L. S., & de Aquino, G. S. (2015). Mobile Application Estimate the Design Phase. In L. A. Maciaszek & J. Filipe (Eds.), *Evaluation of Novel Approaches to Software Engineering* (Vol. 551, pp. 155–167). Springer International Publishing. [https://doi.org/10.1007/978-3-319-27218-4\\_11](https://doi.org/10.1007/978-3-319-27218-4_11)
- Fahmy, S., Ngah, N., Roslina, W., & Fariha, Z. (2012). Evaluating the Quality of Software in e-Book Using the ISO 9126 Model. *International Journal of Control and Automation*, 5, 115–122.

- Fang, J., Li, J., & Wang, R. (2017). Mobile Travel Apps' Adoption: Integrating Perceived Characteristics of Innovation and Software Quality. *Proceedings of the 2017 International Conference on Software and E-Business - ICSEB 2017*, 38–42. <https://doi.org/10.1145/3178212.3178213>
- Fauzia, H., Laksmiwati, Ir. H., & Hendradjaya, B. (2014). A quality model for mobile thick client that utilizes web API. *2014 International Conference on Data and Software Engineering (ICODSE)*, 1–6. <https://doi.org/10.1109/ICODSE.2014.7062681>
- Franch, X., & Carvallo, J. P. (2003). Using quality models in software package selection. *IEEE Software*, 20(1), 34–41. <https://doi.org/10.1109/MS.2003.1159027>
- Franke, D., Kowalewski, S., & Weise, C. (2012). A Mobile Software Quality Model. *2012 12th International Conference on Quality Software*, 154–157. <https://doi.org/10.1109/QSIC.2012.49>
- Franke, D., & Weise, C. (2011). Providing a Software Quality Framework for Testing of Mobile Applications. *2011 Fourth IEEE International Conference on Software Testing, Verification and Validation*, 431–434. <https://doi.org/10.1109/ICST.2011.18>
- Gafni, R. (2009). *Usability Issues in Mobile-Wireless Information Systems*. InSITE 2009: Informing Science + IT Education Conference. <https://doi.org/10.28945/3383>
- Grano, G., Di Sorbo, A., Mercaldo, F., Visaggio, C. A., Canfora, G., & Panichella, S. (2017). Android apps and user feedback: A dataset for software evolution and quality improvement. *Proceedings of the 2nd ACM SIGSOFT International Workshop on App Market Analytics - WAMA 2017*, 8–11. <https://doi.org/10.1145/3121264.3121266>
- Gronli, T.-M., & Ghinea, G. (2016). Meeting Quality Standards for Mobile Application Development in Businesses: A Framework for Cross-Platform Testing. *2016 49th Hawaii International Conference on System Sciences (HICSS)*, 5711–5720. <https://doi.org/10.1109/HICSS.2016.706>
- Guerra, A., & Colombo, R. (2009). Tecnologia de qualidade de Produto de Software. *Tecnologia de Qualidade de Produto de Software*. PBQP-Software, Ministério da Ciência e Tecnologia.
- Haigh, M. (2010). Software quality, non-functional software requirements and IT-business alignment. *Software Quality Journal*, 18(3), 361–385. <https://doi.org/10.1007/s11219-010-9098-3>



- Harrison, R., Flood, D., & Duce, D. (2013). Usability of mobile applications: Literature review and rationale for a new usability model. *Journal of Interaction Science*, 1(1), 1. <https://doi.org/10.1186/2194-0827-1-1>
- Hecht, G., Moha, N., & Rouvoy, R. (2016). An empirical study of the performance impacts of Android code smells. *Proceedings of the International Workshop on Mobile Software Engineering and Systems - MOBILESoft '16*, 59–69. <https://doi.org/10.1145/2897073.2897100>
- Hecht, G., Rouvoy, R., Moha, N., & Duchien, L. (2015). Detecting Antipatterns in Android Apps. *2015 2nd ACM International Conference on Mobile Software Engineering and Systems*, 148–149. <https://doi.org/10.1109/MobileSoft.2015.38>
- Hess, S., Kiefer, F., & Carbon, R. (2012). Quality by Construction Through mConcAppt: Towards Using UI-construction as Driver for High Quality Mobile App Engineering. *2012 Eighth International Conference on the Quality of Information and Communications Technology*, 313–318. <https://doi.org/10.1109/QUATIC.2012.48>
- Hilwa, W., & Samidi. (2014). Prototype mobile Knowledge Management System (KMS) for Islamic banking with Tiwana's framework on university: Case study STEI SEBI. *2014 International Conference on Cyber and IT Service Management (CITSM)*, 83–88. <https://doi.org/10.1109/CITSM.2014.7042181>
- Holl, K., & Elberzhager, F. (2014). A Mobile-Specific Failure Classification and Its Usage to Focus Quality Assurance. *2014 40th EUROMICRO Conference on Software Engineering and Advanced Applications*, 385–388. <https://doi.org/10.1109/SEAA.2014.19>
- Holl, K., & Vieira, V. (2015). Focused Quality Assurance of Mobile Applications: Evaluation of a Failure Pattern Classification. *2015 41st Euromicro Conference on Software Engineering and Advanced Applications*, 349–356. <https://doi.org/10.1109/SEAA.2015.36>
- Holzinger, A., Treitler, P., & Slany, W. (2012). Making Apps Useable on Multiple Different Mobile Platforms: On Interoperability for Business Application Development on Smartphones. In G. Quirchmayr, J. Basl, I. You, L. Xu, & E. Weippl (Eds.), *Multidisciplinary Research and Practice for Information Systems* (Vol. 7465, pp. 176–189). Springer Berlin Heidelberg. [https://doi.org/10.1007/978-3-642-32498-7\\_14](https://doi.org/10.1007/978-3-642-32498-7_14)

- Huang, K.-Y. (2009). Challenges in Human-Computer Interaction Design for Mobile Devices. In Ao, SI and Douglas, C and Grundfest, WS and Burgstone, J (Ed.), *WCECS 2009: WORLD CONGRESS ON ENGINEERING AND COMPUTER SCIENCE, VOLS I AND II* (pp. 236–241). INT ASSOC ENGINEERS-IAENG.
- Hussain, A., & Kutar, M. (2009). *Usability Metric Framework for Mobile Phone Application*.
- Hyun, J. L., & Soo, D. K. (2013). A model of quality-in-use for service-based mobile ecosystem. *2013 1st International Workshop on the Engineering of Mobile-Enabled Systems (MOBS)*, 13–18. <https://doi.org/10.1109/MOBS.2013.6614217>
- Idri, A., Bachiri, M., & Fernández-Alemán, J. L. (2016). A Framework for Evaluating the Software Product Quality of Pregnancy Monitoring Mobile Personal Health Records. *Journal of Medical Systems*, 40(3), 50. <https://doi.org/10.1007/s10916-015-0415-z>
- Idri, A., Bachiri, M., Fernandez-Aleman, J. L., & Toval, A. (2016). Experiment design of free pregnancy monitoring mobile personal health records quality evaluation. *2016 IEEE 18th International Conference on E-Health Networking, Applications and Services (Healthcom)*, 1–6. <https://doi.org/10.1109/HealthCom.2016.7749501>
- Idri, A., Bachiri, M., Fernandez-Aleman, J. L., & Toval, A. (2017). ISO/IEC 25010 Based Evaluation of Free Mobile Personal Health Records for Pregnancy Monitoring. *2017 IEEE 41st Annual Computer Software and Applications Conference (COMPSAC)*, 262–267. <https://doi.org/10.1109/COMPSAC.2017.159>
- ISO/IEC. (2001). *Software engineering – Product quality – Part 1: Quality model (ISO/IEC 9126-1:2001)*.
- ISO/IEC. (2011). *Systems and software engineering—Systems and software Quality Requirements and Evaluation (SQuaRE)—System and software quality models. (ISO/IEC 25010:2011)*.
- ISO/IEC. (2014). *Software engineering—Systems and software Quality Requirements and Evaluation (SQuaRE)—Requirements for quality of Ready to Use Software Product (RUSP) and instructions for testing (ISO/IEC 25051:2014)*.
- ISO/IEC. (2015). *Information technology—Process assessment—Process measurement framework for assessment of process capability ISO/IEC 33020:2015 [ISO/IEC 33020:2015]*.

- Kabir, M. A., Han, J., Colman, A., Aljohani, N. R., Basher, M., Xing, Z., & Lin, S.-W. (2016). Engineering Socially-Aware Systems and Applications. *2016 21st International Conference on Engineering of Complex Computer Systems (ICECCS)*, 81–90. <https://doi.org/10.1109/ICECCS.2016.019>
- Khalid, H., Nagappan, M., & Hassan, A. E. (2016). Examining the Relationship between FindBugs Warnings and App Ratings. *IEEE Software*, 33(4), 34–39. <https://doi.org/10.1109/MS.2015.29>
- Khalid, H., Shihab, E., Nagappan, M., & Hassan, A. E. (2015). What Do Mobile App Users Complain About? *IEEE Software*, 32(3), 70–77. <https://doi.org/10.1109/MS.2014.50>
- Kim, H.-W., Lee, H. L., & Son, J. E. (2011). *An Exploratory Study on the Determinants of Smartphone App Purchase*. 10.
- Kim, M., Park, J. H., & Lee, N. Y. (2017). A quality model for IoT service. In *Lecture Notes in Electrical Engineering* (pp. 497–504). [https://doi.org/10.1007/978-981-10-3023-9\\_77](https://doi.org/10.1007/978-981-10-3023-9_77)
- Kitchenham, B., & Charters, S. (2007). *Guidelines for performing systematic literature reviews in software engineering* (p. 57p.) [Technical Report EBSE 2007-001]. Keele University and Durham University.
- La, H. J., Ho Joong Lee, & Kim, S. D. (2011). An efficiency-centric design methodology for mobile application architectures. *2011 IEEE 7th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, 272–279. <https://doi.org/10.1109/WiMOB.2011.6085388>
- Lai, I. K. W. (2015). Traveler Acceptance of an App-Based Mobile Tour Guide. *Journal of Hospitality & Tourism Research*, 39(3), 401–432. <https://doi.org/10.1177/1096348013491596>
- Lew, P., & Olsina, L. (2013). Relating User Experience with MobileApp Quality Evaluation and Design. In Q. Z. Sheng & J. Kjeldskov (Eds.), *Current Trends in Web Engineering* (Vol. 8295, pp. 253–268). Springer International Publishing. [https://doi.org/10.1007/978-3-319-04244-2\\_23](https://doi.org/10.1007/978-3-319-04244-2_23)
- Liao, G.-Y., Chien, Y.-T., Chen, Y.-J., Hsiung, H.-F., Chen, H.-J., Hsieh, M.-H., & Wu, W.-J. (2017). What to Build for Middle-Agers to Come? Attractive and Necessary Functions of Exercise-Promotion Mobile Phone Apps: A Cross-Sectional Study. *JMIR MHealth and UHealth*, 5(5), e65. <https://doi.org/10.2196/mhealth.6233>

- Lim, S. L., Bentley, P. J., Kanakam, N., Ishikawa, F., & Honiden, S. (2015). Investigating Country Differences in Mobile App User Behavior and Challenges for Software Engineering. *IEEE Transactions on Software Engineering*, 41(1), 40–64. <https://doi.org/10.1109/TSE.2014.2360674>
- Linares-Vásquez, M., Klock, S., McMillan, C., Sabané, A., Poshyvanyk, D., & Guéhéneuc, Y.-G. (2014). Domain matters: Bringing further evidence of the relationships among anti-patterns, application domains, and quality-related metrics in Java mobile apps. *Proceedings of the 22nd International Conference on Program Comprehension - ICPC 2014*, 232–243. <https://doi.org/10.1145/2597008.2597144>
- Liu, Z., Hu, Y., & Cai, L. (2014). Software Quality Testing Model for Mobile Application. In I. Awan, M. Younas, X. Franch, & C. Quer (Eds.), *Mobile Web Information Systems* (Vol. 8640, pp. 192–204). Springer International Publishing. [https://doi.org/10.1007/978-3-319-10359-4\\_16](https://doi.org/10.1007/978-3-319-10359-4_16)
- Marinho, E. H., & Resende, R. F. (2012). Quality Factors in Development Best Practices for Mobile Applications. In B. Murgante, O. Gervasi, S. Misra, N. Nedjah, A. M. A. C. Rocha, D. Taniar, & B. O. Apduhan (Eds.), *Computational Science and Its Applications – ICCSA 2012* (Vol. 7336, pp. 632–645). Springer Berlin Heidelberg. [https://doi.org/10.1007/978-3-642-31128-4\\_47](https://doi.org/10.1007/978-3-642-31128-4_47)
- McCall, J. (1977). *Factors in Software Quality: Preliminary Handbook on Software Quality for an Acquisition Manager* (Vols. 1–3). General Electric. <http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA049055>
- Mohsin, A., Naqvi, S. I. R., Khan, A. U., Naeem, T., & AsadUllah, M. A. (2017). A comprehensive framework to quantify fault tolerance metrics of web centric mobile applications. *2017 International Conference on Communication Technologies (ComTech)*, 65–71. <https://doi.org/10.1109/COMTECH.2017.8065752>
- Moumane, K., & Idri, A. (2017). Software quality in mobile environments: A comparative study. *2017 4th International Conference on Control, Decision and Information Technologies (CoDIT)*, 1123–1128. <https://doi.org/10.1109/CoDIT.2017.8102750>

- Moumane, K., Idri, A., & Abran, A. (2016). Usability evaluation of mobile applications using ISO 9241 and ISO 25062 standards. *SpringerPlus*, 5(1), 548. <https://doi.org/10.1186/s40064-016-2171-z>
- Naab, M., Braun, S., Lenhart, T., Hess, S., Eitel, A., Magin, D., Carbon, R., & Kiefer, F. (2015). Why Data Needs more Attention in Architecture Design—Experiences from Prototyping a Large-Scale Mobile App Ecosystem. *2015 12th Working IEEE/IFIP Conference on Software Architecture*, 75–84. <https://doi.org/10.1109/WICSA.2015.13>
- Nayebi, F., Desharnais, J.-M., & Abran, A. (2012). The state of the art of mobile application usability evaluation. *2012 25th IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)*, 1–4. <https://doi.org/10.1109/CCECE.2012.6334930>
- Olsina, L., & Lew, P. (2017). Specifying mobileapp quality characteristics that may influence trust. *Proceedings of the 13th Central & Eastern European Software Engineering Conference in Russia on - CEE-SECR '17*, 1–9. <https://doi.org/10.1145/3166094.3166097>
- Orru, M., Porru, S., Tonelli, R., & Marchesi, M. (2015). A Preliminary Study on Mobile Apps Call Graphs through a Complex Network Approach. *2015 11th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS)*, 482–488. <https://doi.org/10.1109/SITIS.2015.95>
- Pai, M., McCulloch, M., Gorman, J. D., Pai, N., Enanoria, W., Kennedy, G., Tharyan, P., & Colford, J. M. (2004). Systematic reviews and meta-analyses: An illustrated, step-by-step guide. *The National Medical Journal of India*, 17(2), 86–95.
- Peischl, B., Ferk, M., & Holzinger, A. (2015). The fine art of user-centered software development. *Software Quality Journal*, 23(3), 509–536. <https://doi.org/10.1007/s11219-014-9239-1>
- Petersen, K., Vakkalanka, S., & Kuzniarz, L. (2015). Guidelines for conducting systematic mapping studies in software engineering: An update. *Information and Software Technology*, 64, 1–18. <https://doi.org/10.1016/j.infsof.2015.03.007>
- Potena, P. (2013). Optimization of adaptation plans for a service-oriented architecture with cost, reliability, availability and performance tradeoff. *Journal of Systems and Software*, 86(3), 624–648. <https://doi.org/10.1016/j.jss.2012.10.929>
- Pretel, I., & Lago, A. B. (2012). Mobile-Human Interaction Monitoring System. In J. Del Ser, E. A. Jorswieck, J. Miguez, M. Matinmikko, D. P. Palomar, S. Salcedo-Sanz,

- & S. Gil-Lopez (Eds.), *Mobile Lightweight Wireless Systems* (Vol. 81, pp. 198–205). Springer Berlin Heidelberg. [https://doi.org/10.1007/978-3-642-29479-2\\_15](https://doi.org/10.1007/978-3-642-29479-2_15)
- Rao, L. M., Suma, V., & Shubhamangala, B. R. (2012). Impact analysis of volatility and security on requirement defects during software development process. *International Conference on Software Engineering and Mobile Application Modelling and Development (ICSEMA 2012)*, 12–12. <https://doi.org/10.1049/ic.2012.0145>
- Ribeiro, M. I. C., & Dias-Neto, A. C. (2017). Company Health in Mobile Software Ecosystem (MSECO): Research Perspectives and Challenges. *2017 IEEE/ACM Joint 5th International Workshop on Software Engineering for Systems-of-Systems and 11th Workshop on Distributed Software Development, Software Ecosystems and Systems-of-Systems (JSOS)*, 74–75. <https://doi.org/10.1109/JSOS.2017.3>
- Ricciardi, P., Delaitre, P., Lavandier, C., Torchia, F., & Aumond, P. (2015). Sound quality indicators for urban places in Paris cross-validated by Milan data. *The Journal of the Acoustical Society of America*, 138(4), 2337–2348. <https://doi.org/10.1121/1.4929747>
- Rocha, A., Travassos, G., Santos, G., & Reinehr, S. (2017). *QPS-Modelo para Avaliação da Qualidade de Produtos de Software: Resultados Iniciais*. In Proceedings of the Brazilian Software Quality Symposium (SBQS'17).
- Rodríguez, M., Pedreira, O., & Fernández, C. (2015). Certificación de la Mantenibilidad del Producto Software: Un Caso Práctico. *Revista Latinoamericana de Ingeniería de Software*, 3, 127. <https://doi.org/10.18294/relais.2015.127-134>
- Rohil, M. K., & Gupta, N. K. (2012). Software quality measurement for reusability. *International Conference on Software Engineering and Mobile Application Modelling and Development (ICSEMA 2012)*, 24–24. <https://doi.org/10.1049/ic.2012.0157>
- Ryan, C., & Rossi, P. (2005). Software, Performance and Resource Utilisation Metrics for Context-Aware Mobile Applications. *11th IEEE International Software Metrics Symposium (METRICS'05)*, 12–12. <https://doi.org/10.1109/METRICS.2005.44>
- Savio, N., & Braiterman, J. (2007). Design Sketch: The Context of Mobile Interaction. *International Journal of Mobile Marketing*, 2.

- Seshasayee, B., Nathuji, R., & Schwan, K. (2007). Energy-Aware Mobile Service Overlays: Cooperative Dynamic Power Management in Distributed Mobile Systems. *Fourth International Conference on Autonomic Computing (ICAC'07)*, 6–6. <https://doi.org/10.1109/ICAC.2007.14>
- Soad, G. W., Filho, N. F. D., & Barbosa, E. F. (2016). Quality evaluation of mobile learning applications. *2016 IEEE Frontiers in Education Conference (FIE)*, 1–8. <https://doi.org/10.1109/FIE.2016.7757540>
- Spriestersbach, A., & Springer, T. (2004). Quality Attributes in Mobile Web Application Development. In F. Bomarius & H. Iida (Eds.), *Product Focused Software Process Improvement* (Vol. 3009, pp. 120–130). Springer Berlin Heidelberg. [https://doi.org/10.1007/978-3-540-24659-6\\_9](https://doi.org/10.1007/978-3-540-24659-6_9)
- Syer, M. D., Nagappan, M., Adams, B., & Hassan, A. E. (2015). Studying the relationship between source code quality and mobile platform dependence. *Software Quality Journal*, 23(3), 485–508. <https://doi.org/10.1007/s11219-014-9238-2>
- Trienekens, J. J. M., Kusters, R. J., & Brussel, D. C. (2010). Quality specification and metrication, results from a case-study in a mission-critical software domain. *Software Quality Journal*, 18(4), 469–490. <https://doi.org/10.1007/s11219-010-9101-z>
- Wasserman, A. I. (2010). Software engineering issues for mobile application development. *Proceedings of the FSE/SDP Workshop on Future of Software Engineering Research - FoSER '10*, 397. <https://doi.org/10.1145/1882362.1882443>
- Wewers, M. E., & Lowe, N. K. (1990). A critical review of visual analogue scales in the measurement of clinical phenomena. *Research in Nursing & Health*, 13(4), 227–236. <https://doi.org/10.1002/nur.4770130405>
- Wohlin, C., Runeson, P., Hst, M., Ohlsson, M. C., Regnell, B., & Wessln, A. (2012). *Experimentation in Software Engineering*. Springer Publishing Company, Incorporated.
- Yildiz, E., Bilgen, S., Tokdemir, G., Cagiltay, N. E., & Erturan, Y. N. (2014). Analysis of B2C Mobile Application Characteristics and Quality Factors Based on ISO 25010 Quality Model. In I. Awan, M. Younas, X. Franch, & C. Quer (Eds.), *Mobile Web Information Systems* (Vol. 8640, pp. 261–274). Springer International Publishing. [https://doi.org/10.1007/978-3-319-10359-4\\_21](https://doi.org/10.1007/978-3-319-10359-4_21)

- Zahra, S., Khalid, A., & Javed, A. (2013). An Efficient and Effective New Generation Objective Quality Model for Mobile Applications. *International Journal of Modern Education and Computer Science*, 5(4), 36–42. <https://doi.org/10.5815/ijmecs.2013.04.05>
- Zernadji, T., Tibermachine, C., Cherif, F., & Zouioueche, A. (2016). Integrating quality requirements in engineering web service orchestrations. *Journal of Systems and Software*, 122, 463–483. <https://doi.org/10.1016/j.jss.2015.11.009>
- Zhenyu Liu, Yun Hu, & Lizhi Cai. (2014). Research on software security and compatibility test for mobile application. *Fourth Edition of the International Conference on the Innovative Computing Technology (INTECH 2014)*, 140–145. <https://doi.org/10.1109/INTECH.2014.6927764>



# Annex A – Extraction Form

Version History				
Date	Version	Description	Author	Reviewer

General Information	
Title	
Authors	
Year of Publication	
Publication Source	
Search Engines	
Citation	
Context Approach	
How the characteristics were defined, selected and evaluated?	
Abstract	

Mobile application	
Concept	N/A

Product quality characteristics in ISO/IEC 25010		
Name	Definition	Comment
N/A	N/A	N/A
Product quality sub-characteristics in ISO/IEC 25010		
Name	Definition	Comment
N/A	N/A	N/A

Quality in use characteristics in ISO/IEC 25010		
Name	Definition	Comment
N/A	N/A	N/A
Quality in use sub-characteristics in ISO/IEC 25010		
Name	Definition	Comment
N/A	N/A	N/A

Quality characteristics not present in ISO/IEC 25010		
Name	Definition	Comment
N/A	N/A	N/A

Measures	Definition	Related characteristic
N/A	N/A	N/A

Characteristic of limitation of mobile devices	
Characteristic/Limitation	Comment
N/A	N/A

Mobile application requirements	
Requirement	Related Characteristic
N/A	N/A

## Annex B – Appraisal Instrument

A avaliação de um produto é feita avaliando-se separadamente as características de qualidade do produto (gerais e específicas de aplicativos) e as características de qualidade em uso. Em cada caso a avaliação se dá em três passos: (i) caracterização das sub características de qualidade (ii) caracterização das características de qualidade e (iii) atribuição de nível QPS.

A caracterização das sub características é feita por consenso entre os membros da equipe de avaliação, atribuindo a cada sub característica um dos seguintes graus:

- T (Totalmente atendido)
- L (Largamente atendido)
- P (Parcialmente atendido)
- N (Não atendido)
- NA (Não é possível avaliar)

A **caracterização das características** obedece às seguintes regras:

- Considerar a caracterização das sub características relacionadas à característica.
  - Atribuir 3 a cada T, 2 a cada L, 1 a cada P e 0 a cada N.
  - Ignorar, neste momento, as sub características avaliadas como “Não é possível Avaliar”
- Calcular a mediana dos valores atribuídos a cada sub característica relacionada a uma determinada característica. Ter em conta que o número de sub características relacionadas às características é variável.
- Caracterizar cada característica, atribuindo:
  - T se a mediana for  $>2,55$  até 3
  - L se a mediana for  $>1,5$  até 2,55
  - P se a mediana for  $>0,45$  até 1,5
  - N se a mediana for  $<0,45$

Após a caracterização de cada característica com T, L, P ou N é realizada a atribuição de nível QPS ao produto.

Um produto é **BRONZE** se:

- A sub característica de qualidade SE1-G foi caracterizada como T e as demais sub características de qualidade gerais foram caracterizadas como L ou T.

Um produto é **PRATA** se:

- A sub característica de qualidade SE1-G foi caracterizada como T e as demais características de qualidade gerais foram caracterizadas como L ou T
- As características de qualidade em uso tiverem sido avaliadas como L ou T.

Um produto é **OURO** se:

- Todas as sub características de qualidade gerais foram caracterizadas como T
- As características de qualidade em uso tiverem sido avaliadas como L ou T.
- As características de qualidade do produto foram caracterizadas como L ou T.

## Instrumento de Avaliação

### Avaliação da Qualidade do Produto

#### 1.1 Caracterização das sub características de qualidade do produto

ID	Sub característica	Questão	Caracterização
US1-G	Operabilidade	As tarefas têm comportamento e aparência consistentes?	
US2-G	Operabilidade	As mensagens fornecidas pelo produto são claras?	
US3-G	Operabilidade	Existe undo ou confirmação para tarefas com consequências significativas?	
US4-G	Estética da interface com o usuário	A interface com o usuário é esteticamente agradável?	
US5-E	Reconhecimento da Adequação	Ao usar pela primeira vez, o aplicativo permite ao usuário perceber se ele é adequado às necessidades?	
US6-E	Facilidade de Aprendizado	É fácil aprender como utilizar as funcionalidades do aplicativo?	
US7-E	Operabilidade	É fácil operar e controlar o aplicativo para utilizar as funcionalidades?	
US8-E	Proteção de Erros do Usuário	O aplicativo evita que o usuário cometa erros?	
US9-E	Acessibilidade	O aplicativo tem adaptações que permitam seu uso por pessoas com limitações auditivas, visuais e motoras?	
US10-E	Navegabilidade	É fácil encontrar a funcionalidade ou a informação que o usuário precisa?	
US11-E	Visibilidade da Interface	O aplicativo tem as telas organizadas de modo que seja fácil e rápido o entendimento do texto?	
US12-E	Formulários Claros e Curtos	O aplicativo contém formulários claros e com opção de ajuda?	
US13-E	Menus Hierárquicos	O aplicativo tem poucos menus, e estes são simples e fáceis de navegar?	

SE1-G	Confidencialidade	Existe controle de acesso, isto é proteção contra acessos não autorizados?	
SE2-E	Confidencialidade	O aplicativo pede autorização para usar suas informações?	
SE3-G	Integridade	Existe prevenção para que dados não sejam corrompidos ou modificados por acessos não autorizados?	
AF1-E	Correção Funcional	O aplicativo faz corretamente o que o usuário espera que ele faça?	
ED1-E	Comportamento com Relação ao Tempo	O aplicativo possui tempo de resposta satisfatório?	
ED2-E	Utilização de Recursos	O aplicativo utiliza satisfatoriamente recursos como memória e bateria?	
COMP1-E	Interoperabilidade	O aplicativo comunica dados satisfatoriamente com outros aplicativos?	
CONF1-E	Tolerância a Falhas	O aplicativo se comporta de forma adequada mesmo em caso de problemas no equipamento ou no software?	
CONF2-E	Persistência de Dados	O aplicativo guarda as informações apropriadas em memória após possíveis pausas?	
CONF3-E	Qualidade da Informação	As informações fornecidas pelo aplicativo são precisas e adequadas às necessidades do usuário?	
PORT1-E	Adaptabilidade	O aplicativo é corretamente adaptado para diferentes plataformas?	
MANU1-G	Analisabilidade	A estrutura para rastreabilidade existe e está completa de forma a poder apoiar a análise do impacto de mudanças?	
MANU2-G	Testabilidade	A estrutura para rastreabilidade existe e está completa de forma a poder apoiar a realização de testes após a mudança?	
MANU3-G	Testabilidade	Há a disponibilidade de casos de teste para realização de testes de regressão após mudanças?	

## 1.2 Caracterização das características de qualidade do produto

Característica de qualidade do produto	Caracterização
Usabilidade	
Segurança	
Adequação Funcional	
Eficiência no Desempenho	
Compatibilidade	
Confiabilidade	
Manutenibilidade	

## Avaliação da Qualidade em Uso

### 2.1 Caracterização das sub características de qualidade em uso

ID	Sub característica	Questão	Caracterização
EFI1-S	Eficiência	Ao usar o aplicativo, o usuário atinge seus objetivos sem se cansar?	
EFE1-S	Efetividade	Ao usar o aplicativo, o usuário atinge seus objetivos com precisão e completude?	
SAT1-S	Utilidade	O aplicativo se mostra útil para o usuário alcançar o objetivo que o fez utilizá-lo?	
SAT2-S	Prazer	O uso do aplicativo é agradável?	
COB1-S	Completeza de Contexto	O aplicativo funciona apropriadamente em vários contextos de uso esperados?	
COB2-S	Flexibilidade	O aplicativo funciona apropriadamente em contextos de uso além dos esperados?	
COM1-S	Comunicação Contínua	O uso do aplicativo não é atrapalhado por problemas como falta de internet e sinal de GPS?	

### 2.2 Caracterização das características de qualidade em uso

Característica de qualidade em uso	Caracterização
<b>Eficiência</b>	
<b>Efetividade</b>	
<b>Satisfação</b>	
<b>Cobertura de Contexto</b>	
<b>Comunicação Contínua</b>	

# Annex C - Appraisal Plan

## Informações do produto

**Nome do Produto:** Aplicativo para Celular do Banco X

**Descrição do produto:** Aplicativo bancário que possibilita o gerenciamento rápido da conta do cliente, através do qual pode-se acessar o extrato das contas corrente e poupança, consultar saldo, ver movimentações de cartão de débito, pagar contas e investir em renda fixa e previdência privada.

Algumas funcionalidades (texto extraído da Play Store):

- Transferências com o Teclado X: você não precisa mais sair de uma conversa e mudar de aplicativo para transferir dinheiro para qualquer banco.
- Monitore seus gastos: acompanhe as movimentações, tanto de saques como de compras no cartão de débito e consulte seu saldo, além de extrato na hora que quiser direto do aplicativo do banco online.
- Controle o cartão de crédito: acompanhe os gastos do seu cartão de crédito, consulte sua fatura atual, a próxima e gerencie seu dinheiro.
- Gerencie seus investimentos: pelo aplicativo do banco, você consegue fazer aplicações e resgates de investimentos da poupança e fundo de investimentos.
- Transferências e crédito na hora: transfira para contas do banco X ou qualquer outro banco e faça recargas do seu celular pré-pago pelo app.

**Versão atual:**

**Data de lançamento da versão atual:**

**Escopo da avaliação:** Avaliação de produto já disponível no mercado.

**Exclusões de processos:** Nenhum processo será avaliado, por não se ter acesso à documentação do produto.

**Patrocinador da avaliação:** Não se aplica

## Cronograma da avaliação

Atividade	Responsável	Data
Elaboração do Plano de Avaliação	Avaliador Líder e Coordenador Local	XX/11/2019
Avaliação do produto	Equipe de Avaliação	XX/11/2019
Envio do Relatório Final de Avaliação	Avaliador Líder	XX/11/2019

**Projetos selecionados para avaliação:** Não se aplica

**Serviços selecionados para avaliação:** Não se aplica

## Equipe de avaliação

Nome	Papel na equipe
	Avaliador Líder
	Membro da Equipe de Avaliação

**Equipe envolvida com processos da dimensão organizacional:** Não se aplica

**Equipe envolvida com os projetos (Dimensão Engenharia de Software):** Não se aplica

**Equipe envolvida com os serviços (Dimensão Serviços):** Não se aplica

## Cronograma das atividades

**Atividades do dia 23/11/2019**

Horário	Atividade	Participantes
	Reunião de abertura	Equipe de Avaliação e Coordenador Local
	Avaliação do aplicativo versão iOS	Equipe de Avaliação
	Deliberação e atribuição de nível QPS	Equipe de Avaliação

## Aprovação do Plano da Avaliação

Papel	Nome	Assinatura	Data
Avaliador Líder			XX/11/2019
Coordenador Local			XX/11/2019

# Annex D – Final Report (Filled)

## Informações do produto

**Nome do Produto:** Aplicativo para Celular do Banco X

**Descrição do produto:** Aplicativo bancário que possibilita o gerenciamento rápido da conta do cliente, através do qual pode-se acessar o extrato das contas corrente e poupança, consultar saldo, ver movimentações de cartão de débito, pagar contas e investir em renda fixa e previdência privada.

Algumas funcionalidades (texto extraído da Play Store):

- Transferências com o Teclado X: você não precisa mais sair de uma conversa e mudar de aplicativo para transferir dinheiro para qualquer banco.
- Monitore seus gastos: acompanhe as movimentações, tanto de saques como de compras no cartão de débito e consulte seu saldo, além de extrato na hora que quiser direto do aplicativo do banco online.
- Controle o cartão de crédito: acompanhe os gastos do seu cartão de crédito, consulte sua fatura atual, a próxima e gerencie seu dinheiro.
- Gerencie seus investimentos: pelo aplicativo do banco, você consegue fazer aplicações e resgates de investimentos da poupança e fundo de investimentos.
- Transferências e crédito na hora: transfira para contas do banco X ou qualquer outro banco e faça recargas do seu celular pré-pago pelo app.

**Versão atual:**

**Data de lançamento da versão atual:**

## Parâmetros de Avaliação

**Escopo da avaliação:** Avaliação de produto já disponível no mercado.

**Exclusões de processos:** Nenhum processo foi avaliado por não se ter acesso à documentação do produto. Não foram avaliadas as dimensões organizacionais, de engenharia de software e de serviços do modelo QPS.

**Patrocinador da avaliação:** Não se aplica



## Equipe de avaliação

Nome	Papel na equipe
	Avaliador Líder
	Membro da Equipe de Avaliação

## Caracterização

Após análise da versão do aplicativo, a equipe de avaliação graduou cada requisito do modelo segundo o seguinte critério: T (Totalmente Atendido), L (Largamente atendido), P (Parcialmente atendido), N (Não atendido) ou NA não avaliado.

## Caracterização das sub características de qualidade do produto

ID	Sub característica	Questão	Caracterização
US1-G	Operabilidade	As tarefas têm comportamento e aparência consistentes?	<b>T</b>
US2-G	Operabilidade	As mensagens fornecidas pelo produto são claras?	<b>T</b>
US3-G	Operabilidade	Existe undo ou confirmação para tarefas com consequências significativas?	<b>T</b>
US4-G	Estética da interface com o usuário	A interface com o usuário é esteticamente agradável?	<b>T</b>
US5-E	Reconhecimento da Adequação	Ao usar pela primeira vez, o aplicativo permite ao usuário perceber se ele é adequado às necessidades?	<b>L</b>
US6-E	Facilidade de Aprendizado	É fácil aprender como utilizar as funcionalidades do aplicativo?	<b>T</b>
US7-E	Operabilidade	É fácil operar e controlar o aplicativo para utilizar as funcionalidades?	<b>T</b>
US8-E	Proteção de Erros do Usuário	O aplicativo evita que o usuário cometa erros?	<b>L</b>
US9-E	Acessibilidade	O aplicativo tem adaptações que permitam seu uso por pessoas com limitações auditivas, visuais e motoras?	<b>N</b>
US10-E	Navegabilidade	É fácil encontrar a funcionalidade ou a informação que o usuário precisa?	<b>L</b>
US11-E	Visibilidade da Interface	O aplicativo tem as telas organizadas de modo que seja fácil e rápido o entendimento do texto?	<b>T</b>
US12-E	Formulários Claros e Curtos	O aplicativo contém formulários claros e com opção de ajuda?	<b>T</b>
US13-E	Menus Hierárquicos	O aplicativo tem poucos menus, e estes são simples e fáceis de navegar?	<b>L</b>
SE1-G	Confidencialidade	Existe controle de acesso, isto é proteção contra acessos não autorizados?	<b>T</b>
SE2-E	Confidencialidade	O aplicativo pede autorização para usar suas informações?	<b>T</b>

SE3-G	Integridade	Existe prevenção para que dados não sejam corrompidos ou modificados por acessos não autorizados?	NA
AF1-E	Correção Funcional	O aplicativo faz corretamente o que o usuário espera que ele faça?	T
ED1-E	Comportamento com Relação ao Tempo	O aplicativo possui tempo de resposta satisfatório?	L
ED2-E	Utilização de Recursos	O aplicativo utiliza satisfatoriamente recursos como memória e bateria?	L
COMP1-E	Interoperabilidade	O aplicativo comunica dados satisfatoriamente com outros aplicativos?	L
CONF1-E	Tolerância a Falhas	O aplicativo se comporta de forma adequada mesmo em caso de problemas no equipamento ou no software?	L
CONF2-E	Persistência de Dados	O aplicativo guarda as informações apropriadas em memória após possíveis pausas?	T
CONF3-E	Qualidade da Informação	As informações fornecidas pelo aplicativo são precisas e adequadas às necessidades do usuário?	T
PORT1-E	Adaptabilidade	O aplicativo é corretamente adaptado para diferentes plataformas?	NA
MANU1-G	Analisabilidade	A estrutura para rastreabilidade existe e está completa de forma a poder apoiar a análise do impacto de mudanças?	NA
MANU2-G	Testabilidade	A estrutura para rastreabilidade existe e está completa de forma a poder apoiar a realização de testes após a mudança?	NA
MANU3-G	Testabilidade	Há a disponibilidade de casos de teste para realização de testes de regressão após mudanças?	NA

### Caracterização das características de qualidade do produto

Característica de qualidade do produto	Caracterização
Usabilidade	T
Segurança	T
Adequação Funcional	T
Eficiência no Desempenho	L
Compatibilidade	L
Confiabilidade	T
Manutenibilidade	NA
Portabilidade	NA

### Avaliação da Qualidade em Uso

#### Caracterização das sub características de qualidade em uso

ID	Sub característica	Questão	Caracterização
EFI1-S	Eficiência	Ao usar o aplicativo, o usuário atinge seus objetivos sem se cansar?	T
EFE1-S	Efetividade	Ao usar o aplicativo, o usuário atinge seus objetivos com precisão e completude?	T
SAT1-S	Utilidade	O aplicativo se mostra útil para o usuário alcançar o objetivo que o fez utilizá-lo?	T

SAT2-S	Prazer	O uso do aplicativo é agradável?	<b>T</b>
COB1-S	Completeza de Contexto	O aplicativo funciona apropriadamente em vários contextos de uso esperados?	<b>NA</b>
COB2-S	Flexibilidade	O aplicativo funciona apropriadamente em contextos de uso além dos esperados?	<b>NA</b>
COM1-S	Comunicação Contínua	O uso do aplicativo não é atrapalhado por problemas como falta de internet e sinal de GPS?	<b>N</b>

### Caracterização das características de qualidade em uso

Característica de qualidade em uso	Caracterização
<b>Eficiência</b>	<b>T</b>
<b>Efetividade</b>	<b>T</b>
<b>Satisfação</b>	<b>T</b>
<b>Cobertura de Contexto</b>	<b>T</b>
<b>Comunicação Contínua</b>	<b>N</b>

### Pontos Fortes

- Fácil de usar.

### Pontos Fracos

- Não está disponível para pessoas com necessidades especiais.
- Não é possível utilizar o aplicativo off-line mesmo para tarefas de agendamento, por exemplo.

### Sugestões para Melhoria

- Diminuir a quantidade de botões para tornar a interface mais simples.
- Avisar por mensagens claras quando o usuário cometer erros, como por exemplo não concluir um pagamento ou uma transferência.
- Exibir resultados de busca mais orientados à necessidade do usuário. A busca traz informações excessivas.
- Gerar um tutorial para usuários iniciantes.
- Tornar o aplicativo menos dependente dos recursos/capacidade do dispositivo. Torná-lo menos pesado para o uso, permitindo que seja usado com outras tarefas concorrentes.

### Resultado Final

Como resultado da avaliação o produto **Aplicativo para Celular do Banco X** versão 6.9.7 iOS obteve o nível **BRONZE** do Modelo QPS na dimensão Qualidade do Produto.

## Annex E - Revised Instrument

A avaliação de um produto é feita avaliando-se separadamente as características de qualidade do produto (gerais e específicas de aplicativos) e as características de qualidade em uso. Em cada caso a avaliação se dá em três passos: (i) caracterização das sub características de qualidade (ii) caracterização das características de qualidade e (iii) atribuição de nível QPS.

A caracterização das sub características é feita por consenso entre os membros da equipe de avaliação, atribuindo a cada sub característica um dos seguintes graus:

- T (Totalmente atendido)
- L (Largamente atendido)
- P (Parcialmente atendido)
- N (Não atendido)
- NA (Não é possível avaliar)

A **caracterização das características** obedece às seguintes regras:

- Considerar a caracterização das sub características relacionadas à característica.
  - Atribuir 3 a cada T, 2 a cada L, 1 a cada P e 0 a cada N.
  - Ignorar, neste momento, as sub características avaliadas como “Não é possível Avaliar”
- Calcular a mediana dos valores atribuídos a cada sub característica relacionada a uma determinada característica. Ter em conta que o número de sub características relacionadas às características é variável.
- Caracterizar cada característica, atribuindo:
  - T se a mediana for  $>2,55$  até 3
  - L se a mediana for  $>1,5$  até 2,55
  - P se a mediana for  $>0,45$  até 1,5
  - N se a mediana for  $<0,45$

Após a caracterização de cada característica com T, L, P ou N é realizada a atribuição de nível QPS ao produto.

Um produto é **BRONZE** se:

- A sub característica de qualidade SE1-G foi caracterizada como T e as demais sub características de qualidade gerais foram caracterizadas como L ou T.

Um produto é **PRATA** se:

- A sub característica de qualidade SE1-G foi caracterizada como T e as demais características de qualidade gerais foram caracterizadas como L ou T
- As características de qualidade em uso tiverem sido avaliadas como L ou T.

Um produto é **OURO** se:

- Todas as sub características de qualidade gerais foram caracterizadas como T
- As características de qualidade em uso tiverem sido avaliadas como L ou T.
- As características de qualidade do produto foram caracterizadas como L ou T.

## **Instrumento de Avaliação**

**Nome do Produto:** Aplicativo para Celular do Banco X

**Descrição do produto:** Aplicativo bancário que possibilita o gerenciamento rápido da conta do cliente, através do qual pode-se acessar o extrato das contas corrente e poupança, consultar saldo, ver movimentações de cartão de débito, pagar contas e investir em renda fixa e previdência privada.

Algumas funcionalidades (texto extraído da Play Store):

- Transferências com o Teclado X: você não precisa mais sair de uma conversa e mudar de aplicativo para transferir dinheiro para qualquer banco.
- Monitore seus gastos: acompanhe as movimentações, tanto de saques como de compras no cartão de débito e consulte seu saldo, além de extrato na hora que quiser direto do aplicativo do banco online.

- Controle o cartão de crédito: acompanhe os gastos do seu cartão de crédito, consulte sua fatura atual, a próxima e gerencie seu dinheiro.
- Gerencie seus investimentos: pelo aplicativo do banco, você consegue fazer aplicações e resgates de investimentos da poupança e fundo de investimentos.
- Transferências e crédito na hora: transfira para contas do banco X ou qualquer outro banco e faça recargas do seu celular pré-pago pelo app.

**Versão atual:**

## **Avaliação da Qualidade do Produto**

### **1.1 Caracterização das sub características de qualidade do produto**

<b>ID</b>	<b>Sub característica</b>	<b>Questão</b>	<b>Caracteri-zação</b>
US1-G	Operabilidade	As tarefas têm comportamento e aparência consistentes?	
US2-G	Operabilidade	As mensagens fornecidas pelo produto são claras?	
US3-G	Operabilidade	Existe undo ou confirmação para tarefas com consequências significativas?	
US4-G	Estética da interface com o usuário	A interface com o usuário é esteticamente agradável?	
US5-E	Reconhecimento da Adequação	Ao usar pela primeira vez, o aplicativo permite ao usuário perceber se ele é adequado às necessidades?	
US6-E	Facilidade de Aprendizado	É fácil aprender como utilizar as funcionalidades do aplicativo?	
US7-E	Operabilidade	É fácil operar e controlar o aplicativo para utilizar as funcionalidades?	
US8-E	Proteção de Erros do Usuário	O aplicativo evita que o usuário cometa erros?	
US9-E	Acessibilidade	O aplicativo tem adaptações que permitam seu uso por pessoas com limitações auditivas, visuais e motoras?	
US10-E	Navegabilidade	É fácil encontrar a funcionalidade ou a informação que o usuário precisa?	
US11-E	Visibilidade da Interface	O aplicativo tem as telas organizadas de modo que seja fácil e rápido o entendimento do texto?	
US12-E	Formulários Claros e Curtos	O aplicativo contém formulários claros e com opção de ajuda?	
US13-E	Menus Hierárquicos	O aplicativo tem poucos menus, e estes são simples e fáceis de navegar?	
SE1-G	Confidencialidade	Existe controle de acesso, isto é proteção contra acessos não autorizados?	
SE2-E	Confidencialidade	O aplicativo pede autorização para usar suas informações?	

SE3-G	Integridade	Existe prevenção para que dados não sejam corrompidos ou modificados por acessos não autorizados?	
AF1-E	Correção Funcional	O aplicativo faz corretamente o que o usuário espera que ele faça?	
ED1-E	Comportamento com Relação ao Tempo	O aplicativo possui tempo de resposta satisfatório?	
ED2-E	Utilização de Recursos	O aplicativo utiliza satisfatoriamente recursos como memória e bateria?	
COMP1-E	Interoperabilidade	O aplicativo comunica dados satisfatoriamente com outros aplicativos?	
CONF1-E	Tolerância a Falhas	O aplicativo se comporta de forma adequada mesmo em caso de problemas no equipamento ou no software? <b>Exemplo: Em caso de smartphone ou tablete com defeito, ou em caso de sistema operacional Android corrompido, estes não afetam o uso do aplicativo.</b>	
CONF2-E	Persistência de Dados	O aplicativo guarda as informações apropriadas em memória após possíveis pausas?	
CONF3-E	Qualidade da Informação	As informações fornecidas pelo aplicativo são precisas e adequadas às necessidades do usuário?	
PORT1-E	Adaptabilidade	O aplicativo é corretamente adaptado para diferentes plataformas?	
MANU1-G	Analisabilidade	A estrutura para rastreabilidade existe e está completa de forma a poder apoiar a análise do impacto de mudanças?	
MANU2-G	Testabilidade	A estrutura para rastreabilidade existe e está completa de forma a poder apoiar a realização de testes após a mudança?	
MANU3-G	Testabilidade	Há a disponibilidade de casos de teste para realização de testes de regressão após mudanças?	

## 1.2 Caracterização das características de qualidade do produto

Característica de qualidade do produto	Caracterização
Usabilidade	
Segurança	
Adequação Funcional	
Eficiência no Desempenho	
Compatibilidade	
Confiabilidade	
Manutenibilidade	

## Avaliação da Qualidade em Uso

### 2.1 Caracterização das sub características de qualidade em uso

ID	Sub característica	Questão	Caracterização
EFI1-S	Eficiência	Ao usar o aplicativo, o usuário atinge seus objetivos sem se cansar?	
EFE1-S	Efetividade	Ao usar o aplicativo, o usuário atinge seus objetivos com precisão e completude?	

SAT1-S	Utilidade	O aplicativo se mostra útil para o usuário alcançar o objetivo que o fez utilizá-lo?	
SAT2-S	Prazer	O uso do aplicativo é agradável?	
COB1-S	Completeza de Contexto	O aplicativo funciona apropriadamente em vários contextos de uso esperados? <b>Exemplo: Caso o aplicativo seja desenvolvido para uso em áreas abertas, ele abrange todos os possíveis cenários de uso, como falta de atenção do usuário, internet fraca ou ausente, movimentação do dispositivo.</b>	
COB2-S	Flexibilidade	O aplicativo funciona apropriadamente em contextos de uso além dos esperados? <b>Exemplo: Caso o aplicativo seja desenvolvido para uso em áreas abertas, ele também funcionar bem em áreas fechadas, com parâmetros opostos aos especificados.</b>	
COM1-S	Comunicação Contínua	O uso do aplicativo não é atrapalhado por problemas como falta de internet e sinal de GPS?	

## 2.2 Caracterização das características de qualidade em uso

Característica de qualidade em uso	Caracterização
Eficiência	
Efetividade	
Satisfação	
Cobertura de Contexto	
Comunicação Contínua	

**Pontos Fortes:**

**Pontos Fracos:**

**Sugestões para Melhoria:**