Developing a usability method for assessment of M-Commerce systems: a case study at Ericsson

Gabriela Novak

Faculty of Computing
Blekinge Institute of Technology
SE-371 79 Karlskrona Sweden
This thesis is submitted to the Faculty of Computing at Blekinge Institute of Technology in partial fulfillment of the requirements for the degree of Master of Science in Computer Science. The thesis is equivalent to 20 weeks of full time studies.

Contact Information:
Author(s):
Gabriela Novak
E-mail: gabriela.novak@ericsson.com

External advisor:
Shiva Abdi Farzaneh Pour

University advisor:
Lars Lundberg, Ph.D.
Dept. of Computer Science and Engineering

Faculty of Computing
Blekinge Institute of Technology
SE-371 79 Karlskrona, Sweden

Internet : www.bth.se
Phone : +46 455 38 50 00
Fax : +46 455 38 50 57
ABSTRACT

Context. Usability work in software engineering is a measure of quality and it contributes to the overall acceptability of systems. However it is also the most neglected process in the software industry. While there are established guidelines and methods in usability, they are not used in the industry to the full extent.

Objectives. In this case study I examine the level of usability and usability issues with the context of use in an M-Commerce solution. Also, I address the distance between development and the current and potential expert users of the wallet platform solution.

Methods. In this exploratory research, a number of article sources such as IEEE Xplore, ACM Digital Library, and Springer Link are used. Studies are selected after reading titles, abstracts and keywords, then chosen if relevant to the subject. The methods used in this study were literature review, case study and experiment.

Results. A usability test was performed on a specific user interface, in order to detect potential usability issues that might have been overlooked in the development cycle. As an experiment, the test was performed with proxy users and verified with the actual users. A recommendation list based on the test results was produced for possible improvements in the interface.

Conclusions. As a conclusion, a modified usability testing method is proposed. Also I conclude that performing usability testing based on the context of use and the ISO 9241-11 standard, brings value to Ericsson’s current and potential customers as well as to the solution itself. The results from the two groups used for testing were very similar and the proxy user group is thus a good alternative to actual users. With the engagement of a team that is working with the customers dispersed over the world, the context of use can be brought to the development department.

Keywords: usability, usability test, context of use, M-Commerce
ACKNOWLEDGMENTS

I would like to thank my supervisors at BTH and at Ericsson for their guidance, inspiration and enthusiasm that made this thesis possible.

I extend my gratitude to my colleagues at Ericsson, for their participation in the project, their contributions in the usability testing, contextual interviews and overall support with the studies.

This thesis is dedicated to my loving family and friends in Slovenia, who have supported and encouraged me along the way of my studying process and gave me the strength to pursue my dreams.
# CONTENTS

1 INTRODUCTION ........................................................................................................... 5  
  1.1 BACKGROUND ....................................................................................................... 5  
  1.2 AIM AND OBJECTIVES .......................................................................................... 8  
  1.3 RESEARCH QUESTION(S) ....................................................................................... 8  

2 RESEARCH METHODOLOGY ...................................................................................... 9  
  2.1 LITERATURE REVIEW ............................................................................................. 9  
  2.2 CASE STUDY .......................................................................................................... 9  
  2.3 EXPERIMENT ......................................................................................................... 11  

3 USABILITY OVERVIEW ............................................................................................... 12  
  3.1 DEFINITIONS OF USABILITY .................................................................................. 12  
  3.2 WHY USABILITY TESTING? ..................................................................................... 14  
  3.3 USABILITY AND CONTEXT OF USE – ISO 9241-11 STANDARD .......................... 15  
  3.4 USABILITY ENGINEERING ..................................................................................... 17  
  3.5 USABILITY EVALUATION METHODS ...................................................................... 18  
  3.6 THE DISTANCE PROBLEM ..................................................................................... 20  

4 M-COMMERCE AND ERICSSON’S SOLUTION ............................................................... 22  
  4.1 M-COMMERCE ....................................................................................................... 22  
  4.2 ERICSSON WALLET PLATFORM .......................................................................... 23  
  4.3 THE INTERFACE – CUSTOMER CARE GUI ............................................................. 25  
  4.4 USABILITY IN THE INTERFACE ............................................................................. 26  
  4.5 CASE STUDY ......................................................................................................... 26  
  4.6 USABILITY TEST DESIGN ....................................................................................... 26  
    4.6.1 Test preparation ................................................................................................. 28  
    4.6.2 Test participants ............................................................................................... 29  
    4.6.3 Task definition .................................................................................................. 30  
    4.6.4 Test execution ................................................................................................... 31  

5 TEST RESULTS AND ANALYSIS ................................................................................ 32  
  5.1 USABILITY TEST RESULTS – UTC01 – TRANSFER OF FUNDS .............................. 33  
  5.2 USABILITY TEST RESULTS – UTC02 – EDIT ACCOUNT HOLDER INFORMATION ... 34  
  5.3 USABILITY TEST RESULTS – UTC03 – VIEW TRANSACTION HISTORY ............. 34  
  5.4 USABILITY TEST RESULTS – UTC04 – FOUR-EYE-PRINCIPLE ............................ 35  
  5.5 USABILITY TEST OBSERVATION ........................................................................... 35  

6 DISCUSSION ................................................................................................................ 36  
  6.1 PROPOSED USABILITY METHOD EVALUATION – INTERNAL ............................... 36  
    6.1.1 Usability test results – UTC01 – Transfer of funds ........................................... 38  
    6.1.2 Usability test results – UTC02 – Edit account holder information ..................... 38  
    6.1.3 Usability test results – UTC03 – View transaction history ................................. 39  
    6.1.4 Usability test results – UTC04 – Four-eye-principle .......................................... 39  
    6.1.5 Usability test summary ...................................................................................... 39
1 INTRODUCTION

This thesis project targets the expert group of users that work with web based interfaces on their PC, workstations and laptops at the mobile network operator’s offices. The intent with this thesis is to develop a usability method that will serve as an assessment tool for mobile commerce (M-Commerce) systems. With the method the actual context of use will be assessed and feedback from the group of experts gathered. The result will be used at the Ericsson telecommunication company for further development of a system and also to be expanded for more interfaces (mobile phones). With the use of the proposed method, usability issues will be detected in the development phase and the system will be improved accordingly, prior to deployment on a remote location. The long term objective of this type of usability testing and this thesis work is in delivering a usable, highly secure payment system that the customers trust and use with joy.

1.1 Background

After the expansion and success of internet banking systems, a more technologically advanced area is on the rise. With mobile devices and wireless networks the field of M-Commerce is rapidly growing and expanding.

Ericsson has a legacy of developing systems that enable communication and connectivity among people and societies that dates back to the end of 19th century1. With the evolution of communication equipment and software systems intended to enable a networked society, the demands on development and quality are high. More and more systems can be accessed via the World Wide Web or smartphones, and therefore providing good usability is very important.

There are a large number of mobile wallet platforms available which due to lack of security, flexibility and scalability remains un-activated and not used. According to the market research of GSMA Intelligence2, the deployment of mobile wallets is rapidly increasing and will continue in the future. A large amount of these deployments happen in developing countries, where there currently is about 2 billion people possessing a mobile device, but lacking a bank account. However, M-Commerce solutions enable payment services such as cash handling, savings, loans, insurances, but also services like promotions, loyalty programs, tickets, receipts, etc.

Ericsson is developing a mobile wallet platform3 which is secure, highly flexible and will in the future connect more and more societies. The system is developed in Karlskrona and it is in the beginning stages of being rolled-out in the market. The software system – Ericsson Wallet Platform (EWP) is designed with scalability in mind allowing it to grow as the subscriber base grows. It can be installed as a standalone solution or integrated into an existing mobile network with the possibility to be connected to other business support systems such as Ericsson Charging System.

1 http://www.ericsson.com/thecompany/company_facts/history (Visited April 2014)
2 https://gsmaintelligence.com/analysis/2014/02/mobile-platform-wars/420/ (Visited April 2014)
3 http://www.ericsson.com/ourportfolio/products/m-commerce-1 (Visited April 2014)
EWP consists of core modules such as:

- Mobile Money Manager – heart of the system, used for management of all financial transactions and accounts
- Access Module – acts as mediator between the system and other business systems
- Database – data repository of the system
- Customer Care & Reporting Server – server provides all functionality for customer access and administrative work to be managed by expert users. It contains graphical user interfaces for performing tasks related to customer care.

EWP is to be deployed at a mobile network operator’s business which is currently targeting developing countries. The interface to EWP is used by an expert group of people that is authorized to perform registrations, account management, adjustments and similar; these tasks represent the context of use. They perform registration for agents stationed in the field. With their mobile devices they provide services to the end-users. This interface is a very important part of the system; findings from this case will enhance understanding of the customers’ requirements and is therefore the object of usability testing for the purpose of this thesis work.

Due to time limitations, lack of allocated resources and difficulties of launching a new product, usability testing of the EWP interface has not yet been conducted at Ericsson. The established testing strategy during development addresses functionality, robustness, and include traffic simulation of the entire system, whilst the user interfaces have only been tested with an exploratory method. In order to ensure the customers get a usable, high quality product, usability testing shall also be performed and results of the testing considered in further development of the system.

In literature, usability can be assessed with use of quality measurements and can in general serve as a tool to:

- Track progress between releases
- Assessing the competitive position
- Support decision making before launch of a solution
- Creating bonus plans for design managers and higher-level executives

The important standard that serves as a baseline in industrial usability testing is the ISO 9241-11, which defines usability as “the effectiveness, efficiency and satisfaction with which specified users achieve specified goals in particular environments”[1]. The previously mentioned definitions are combined with the use of aspects of the situation, which is referred to as a product’s context of use. The standard also provides an explanation on how measures of user performance and satisfaction can be used to assess any component of the working system and how it affects the whole work system in use.

4 http://www.nngroup.com/articles/usability-metrics/ (Visited April 2014)
Usability guidelines and standards have been developed over the years, and those methods are studied and applied in the industry. However, the usability guidelines provided by the Nielsen Group\(^5\) target mass market web based interfaces and are therefore in need of adjustment and possible expansion for the context of use in the M-Commerce area. As previously mentioned, this is a fast developing area with the aim of connecting people and enabling financial transactions through mobile and wireless networks. In this specific area, established general usability guidelines and the use of connected measures cannot be easily applied. General usability guidelines that also include expectation, page view/clicks, and conversion\(^6\) are intended for a more general group of users and should therefore not be considered as a ready to hand measure of usability in an M-Commerce system.

The expert group in this project that is using the interfaces should be considered – their work processes, tasks that need to be completed on a daily basis to make their work and work through the system interfaces satisfactory. The expert group has a financial background and also technical background and can therefore be considered as skilled users.

The usability methods that are established have remained unchanged for years and can be taken as general guidelines, but in the current expansion of M-Commerce systems, the context of use prevails as highly important – however difficult to understand and grasp in the development phase. In the example of this case study, the long distance between the site of development and the context of use is crucial problem to consider. The first deployment is ongoing in different countries involving different expert groups in Africa and will in the future also expand to countries on other continents. The EWP is highly flexible and can be adjusted to specific customer needs, which the solution architects (SA-team) are collecting as requirements at customer sites. Such information gathering is capturing contextual information and provides a sound basis for later evaluation activities \([2]\). Those system requirements are important in the overall system configuration and will be used as input for usability testing. They provide the context of use view of usability – which depends on the nature of the user, product, task and environment \([3]\).

When assessing a developed system, the context of use is often overlooked. Systems or products are simply divided into those that are usable and the ones that are not. But a system or product cannot be described as usable without taking the context of the product into account \([4]\). The SA-team from the integration department in Karlskrona is a part of this usability test in order to verify the developed assessment method and its application on a configured system. We will therefore refer to the SA-team as proxy users. Their insight and customer understanding is a needed contribution and brings the actual users’ needs to the EWP development in Karlskrona.

Various studies within banking and payments sector have shown the importance of context of use \([5]\). An important contribution to the field of usability is done in the health sector, where the authors approached the problem of using software solutions in a hospital research center in Norway \([6]\). They have illustrated with three cases, addressing the user group, physical environment and scenario. The health sector has specific users, which they take into account while performing usability testing. Their suggestion is in using the ISO definition as a powerful guiding principle but also considering the context of use.


1.2 Aim and objectives

The aim of this thesis is to study the problem of distance between the development and the context of use (financial M-Commerce system) with an extended usability method. This case of investigation will cover the daily operational tasks of the expert group of users and their work with the interfaces of a projected support system.

Objectives of this thesis work:

- Insight to usability testing in the industry
- Measuring the level and importance of usability in further development by taking Ericsson’s solution and its web based interface as the object of study
- Simulating the use environment in Karlskrona with resources that are part of the service delivery department, involving SA-team interacting with real users at customer sites gathering requirements for system adjustments and development
- Providing results of the usability testing
- A method developed would be a step in usability testing for future releases of the solution and also for mobile interfaces
- Recommendations for usability testing of future development of the solution
- Generalization of the method with the possibility of reuse for other stakeholders besides Ericsson.

1.3 Research question(s)

RQ1: How to develop a new M-Commerce usability method which addresses the problem of distance between development and context of use?

RQ2: Can proxy users serve as an alternative in usability testing when real users are not easily accessible?
2 RESEARCH METHODOLOGY

This chapter describes the research methodology used in this study. A case study methodology was applied, after a study of theoretical knowledge in usability testing and applications of such testing in practice. As a quantitative part of the case study and data collection, a usability test with the focus on context of use was performed; qualitative data was gathered from interviews with the SA-team and the colleagues at Ericsson. Object of the usability test was a web based interface to EWP system, which is in use at customer site.

2.1 Literature review

First steps of this study consisted of a detailed review in the usability testing and the context of use. Material studied was mainly acquired through BTH library system\(^7\), Libris\(^8\) and Google Scholar\(^9\) and it is referenced in this thesis. Information retrieval consisted of using keywords such as M-Commerce, Usability Engineering, Usability Testing Methods, Context of Use, Real user Testing, Surrogate Users, Proxy Users and many more. Relevant published material from research communities, different literature, conference proceedings and journals were studied and a research technique called snowball sampling was used. Snowball sampling uses an initial data set, which is the basis for further allocation of additional data \([39]\). Initial relevant published material with its references served as a starting point and further information retrieval was conducted. Results of the literature review and key findings are presented in chapter 3.

2.2 Case study

Definition of a case study is by Yin [8]:

“A case study is an empirical inquiry that investigates a contemporary phenomenon in depth and with its real-life context, especially when the boundary between phenomenon and context is not clearly evident”.

Data collection in a case study consists of use of different sources of information that can be used. Importance of using several data sources is crucial in order to limit the effects of one interpretation of single data source. Conclusions are drawn by analyzing differences between data sources [9].

Different types of data collection beneficial for case studies in software engineering typically involve and were applied for the purpose of this case study:

- Interviews. Interview questions are based on the topic of interest in the case study. Commonly one interview is conducted with every single subject, but

\(^7\) http://www.bth.se/eng/library/ (Visited March – June 2014)

\(^8\) http://libris.kb.se/ (Visited March – June 2014)

\(^9\) http://scholar.google.se/ (Visited March – June 2014)
group interviews are also possible. Questions can be open – allowing a broad range of answers or closed – offering a limited set of alternative answers [9]. Interviews were performed with the SA-team which is gathering requirements for solution adjustments and are in contact with the expert users of the interface that is to be tested. This requirement gathering is a process that involves understanding the customer needs (the expert users), their expectations of the system and the intended use. In software engineering, such requirements are referred to as functional specification\(^\text{10}\). It also includes financial regulations for specific countries that the system must be compliant to. Output of this process is a solution description with all customer requirements and input to development of the system.

- **Observations.** Observations are usually conducted in order to investigate how a certain task is conducted by the test group. Approaches include video recording of the test group and later analyzing the recording; applying a “think aloud” protocol where the researcher is during task execution continuously asking questions like “What is your strategy?” and “What are you thinking?” to remind the test participants to think aloud [9]. A group of 5 people was selected among the solution architects to perform the test. In the Nielsen Norman Group usability research of the optimal number of testers is five, with adding more users to the test group, the same findings are observed and not much new is discovered\(^\text{11}\). They will execute the use cases and their performance will be assessed and recorded for analysis in combination with the think aloud method. Method is considered to be the number one usability tool according to Nielsen Norman Group\(^\text{12}\).

- **Archival data.** This type of data refers to different types of data in the form of documents from different development phases, meeting minutes, financial records and similar [9]. Solution description of the EWP provided by the SA-team as an output of requirement analysis was studied in great detail. Test cases that cover the system’s intended use and context of use were based on this document.

- **Measurements.** Previously mentioned data collection has its focus on qualitative data. Never the less, quantitative data is also important in a case study. Software measurement consists of representing software entities such as processes, products and resources in quantitative numbers [9]. Usability test conducted will cover 4 context specific use cases that are used on a daily basis by the expert users. Design of the test will be done in collaboration with the SA-team that has an understanding of the customer and the system’s intended use. Usability will be tested according to the ISO standard [1].

- **Data analysis is important in case studies and it uses two different approaches. Qualitative data is typically analyzed with the use of descriptive statistics, correlation analysis, development of predictive models and hypothesis testing. Case study research is flexible research method and therefore quantitative and qualitative analysis methods are used. The basic objective of the analysis is to derive conclusions from the data. Information from each step of the study and decisions taken by the researcher must be presented. Analysis must be carried**

\(^{10}\) [http://searchsoftwarequality.techtarget.com/definition/requirements-analysis](http://searchsoftwarequality.techtarget.com/definition/requirements-analysis), (Visited May 2014)


out with the data collection due to the nature of the flexible approach and new insights are found during the analysis [9]. Result of the usability test – qualitative and quantitative measurements are presented in a separate chapter and conclusions are made based on the results. Results will also be used as input for future development and possible improvements of the EWP system.

Conclusions must be reported and set into a context of implications, whether to form theories or to extend the existing knowledge. Conclusions can be drawn without statistics, and they may be interpreted and related to other cases [9]. Usability method that is an object of this case study undergoes two stages of validation:

- Internal evaluation will be performed by the User Experience department at Ericsson. Results of the test and evaluation of the proposed method will in the future serve as an input to further work within development of the EWP system.

- External evaluation will be performed as evaluation of the method beyond this case study. Method will be compared to relevant work published, with its similarities and distinctions. Potential benefits and challenges will be outlined and presented in the conclusions chapter.

2.3 Experiment

Definition of an experiment is by Robson [35]:

"Measuring the effects of manipulating one variable on another variable”.

Experiments are commonly conducted with a limited scope and often run in a laboratory setting. Subjects are assigned to different treatments; the objective is to manipulate one or more variables and controlling all other variables at fixed levels [7].

For the purpose of this research, the treatment consists of a usability test based on the context of use. Subjects in this experiment are:

- Proxy users – they will perform the first iteration of the test. Group consists of SA-team with specific EWP solution knowledge and customer interaction.

- Actual users – they will perform the second iteration of the test. Group consists of the intended users of the EWP solution

Results (see measurements p. 20) will be collected, analyzed and comparison of both test iterations will be presented. The intent with this experiment is to discover if proxy users can detect similar usability issues as the actual users.

This is an empirical, single case study conducted in order to generate theory and extend the knowledge of the research community with a modified usability method applied in the newly emerging area of M-Commerce and an alternative to real users for usability activities.
3 USABILITY OVERVIEW

This chapter is divided into subchapters that represent the overview of the literature studied, from definitions to application of these in similar research papers.

The importance of usability and usability related work has been researched through different research communities of which the Norman Nielsen Group\(^{13}\) is the most beneficial and continuous. The group specializes in conduction of research, user interface evaluation and reporting findings. Work of the community is available online in the form of published articles in the field and they also provide training in usability activities. The driving force behind the group is prof. Jakob Nielsen, Ph.D.,\(^{14}\) who has over the years published a significant amount of work related to usability engineering.

3.1 Definitions of usability

With the term usability three most commonly used definitions are associated.

According to Nielsen, usability is defined as a process in software development cycle and takes into consideration various factors:

- **Learnability:** The system should be easy to learn, so that the user can rapidly start getting some work done with the system.

- **Efficiency:** The system should be efficient to use, so that once the user has learned the system, a high level of productivity is possible.

- **Memorability:** The system should be easy to remember, so that the causal user is able to return to the system after some period of not having used it, without having to learn everything all over again.

- **Errors:** The system should have a low error rate, so that the users make few errors during the use of the system, and so if they do make errors, they can easily recover from them. Further, catastrophic errors must not occur.

- **Satisfaction:** The system should be pleasant to use, so that the users are subjectively satisfied when using it; they like it \(^{11}\).

With further definition of the described factors, usability can be systematically approached, improved and evaluated. Typical measurement process includes a group of test users that uses the system to perform pre-specified set of tasks. The same system can be measured as having different usability characteristics if used by different users for different tasks.

Shneiderman in his work addresses the field of universal usability, and as such he defines usability as having more than 90% of all households as successful users of information and communications services at least once a week. This definition considers factors such as:

\(^{13}\) http://www.nngroup.com/ (Visited May 2014)

\(^{14}\) http://www.nngroup.com/people/jakob-nielsen/ (Visited May 2014)
- Technology variety. Support for a broad range of hardware, software and network access.

- User diversity. Meeting the needs of users with different skills, knowledge, age, gender, handicaps, literacy, etc.

- Gaps in user knowledge. Bridging the gap between what users know and what they need to know [12].

Usability according to the ISO 9241-11 standard is defined as the effectiveness, efficiency and satisfaction with which specified users achieve specified goals in particular environments [1].

Nielsen’s and Shneiderman’s usability definitions target a broader audience in its applications, where the ISO standard focuses on the context of use. Usability in question, with its attributes, for the purpose of this thesis is the ISO standard and its variables are presented in more detail in a separate chapter.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Efficiency</td>
<td>Efficiency</td>
<td>Speed of Performance</td>
</tr>
<tr>
<td></td>
<td>Learnability</td>
<td>Time to Learn</td>
</tr>
<tr>
<td>Effectiveness</td>
<td>Memorability</td>
<td>Retention over time</td>
</tr>
<tr>
<td></td>
<td>Errors</td>
<td>Rate of errors by users</td>
</tr>
<tr>
<td></td>
<td>Safety</td>
<td></td>
</tr>
<tr>
<td>Satisfaction</td>
<td>Satisfaction</td>
<td>Subjective satisfaction</td>
</tr>
</tbody>
</table>

Figure 1: Definitions and variables mapped to one another (source: internet; output from google search)

As mentioned before, important research community within usability testing is the Nielsen Norman Group, which is very active in its research field and their findings are presented in the form of articles published online. Nielsen addresses the durability of usability guidelines in one of his published articles, where he concludes that approximately 90% of usability guidelines from 1986 are still valid, although several of these are in this day less important, because they relate to out dated design elements. Usability guidelines endure because they are dependent on human behavior, which hardly changes – difficulties with use 20 years ago are still present today15.

“Usability guidelines live for a long time, usability methods even longer”, that is due to slow human behavior changes as opposed to the technology evolution16.


3.2 Why usability testing?

Usability studies and testing can usually be done with a small budget and within a short amount of time. Financial payoffs and cost savings from the increased usability commonly become evident after the release of the product [13]. It is however an important part of system development and should not be overlooked.

According to Carter, usability is simple. If we want to know if something is useful, we use it and we find out. Usability testing should not be underestimated just because it is simple [14].

Usability testing consists of measuring the performance of typical users on typical tasks. In a broader sense it relies on a combination of techniques, such as observations, questionnaires, interviews and user testing [37].

Testing with real users is considered to be the most fundamental usability method, since it serves the purpose of direct information about how people use computers, what their exact problems are with a concrete interface being tested. Reliability of user tests possesses a problem because of individual differences between test users. Some users are faster in completing a given task compared to others, and that factor is also important when gathering a group of users. Validity answers the question of whether the usability test in fact measures something of relevance to usability of real products in real use outside of a laboratory [15].

According to Nielsen, usability testing consists of the following activities:

- Test goals and test plans. Clearly defining the purpose of the testing.
- Test users. They should be as representative as possible of the intended users of the system.
- Choosing experimenters. Suitable person to conduct the testing and be in charge of running the test.
- Ethnical aspects. Tests should be conducted with respect to user’s emotions and well-being.
- Test tasks. They should be chosen to be as representative as possible to the uses of the system.
- Performance measurement. Commonly it is measured with a group of test users that perform a set of predefined tasks while collecting time and error data.
- Thinking aloud. Considered as the fundamental and most valuable usability engineering method [15].

Usability has been an area within information technology (IT) systems development and it is growing in importance, mainly due to the fact that the developers mostly work within an abstract world, where foreseeing real-world design problems is at times difficult. Usability
testing helps companies developing systems to discover issues with use and correct them before products emerge onto the market\textsuperscript{17}.

Whether a product is used widely or at work, poor usability inflicts costs on the producer of the product. Hard to use products come in hand with higher support costs, more time and resources are spent on rework and customers tend to be less satisfied. These consequences can often be avoided with application of user center design process and usability engineering during the development of a product [16].

As Cooper states, thoughtful user testing can uncover designer’s incorrect assumptions. Exposing design work to users and then redesigning iteratively is always better than not doing so [38].

With that being said and usability explained with the user in focus and not the product, the ISO standard with its concise definition is presented in the next sub-chapter.

### 3.3 Usability and context of use – ISO 9241-11 standard

The ISO standard addresses designing and evaluating visual display terminals (VDT) for usability, which enables users to achieve goals and meet the needs in a particular context of use. It explains the benefits of user performance and satisfaction. These are measured by the extent, to which the intended goals of use are achieved, the resources that need to be used to achieve the intended goals and the extent to which the user finds the use of the product acceptable [1].

Context of use consists of the users, the tasks, the equipment (hardware, software and materials) and the physical and social environments which may all influence the usability of a product in a work system. Measures of user performance and satisfaction assesses the overall work system, and when a product is the focus of concern, these measures provide information about usability of that product in the particular context of use, provided by the rest of the work system. The effect of changes in other components of the work system can also be measured (amount of user training, improvement of lightning and similar) by user performance and satisfaction [1].

The standard includes guidance on how the usability of a product can be specified and evaluated. It applies for products in general use and also for products designed for specific organizations. It also contains explanation how measures of user performance and satisfaction can be used to measure how any component of a work system affects the whole system in use [1].

Benefits of the use of the standard is in the usability improvement with including features and attributes known to benefit the users in a particular context of use.

\textsuperscript{17} \url{http://newsroom.cisco.com/feature-content?type=webcontent&articleId=1415749} (Visited May 2014)
3.3.1 Usability components

ISO standard forms a framework for specifying usability. Different components of usability and the relationships between them are presented in the figure:

![Usability framework](source: ISO 9241-11)

Figure 2: Usability framework (source: ISO 9241-11)

3.3.1.1 Description of goals

Goals of using the product should be described.

3.3.1.2 Context of use

- Description of users. Characteristics of users that are of importance need to be described. In some cases it may be necessary to define characteristics of different types of users, such as different roles or levels of experience.

- Description of tasks. Characteristics of tasks that need to be completed in order to achieve a goal should be described (frequency or duration of the task). For the evaluation process a set of key tasks will typically be selected to represent the aspects of overall work.

- Description of equipment. Relevant characteristics of the equipment need to be described.

- Description of environments. Relevant characteristics of the physical and social environment need to be described. These include attributes of the wider technical environment, physical environment, ambient environment and the social and cultural environment.

3.3.1.3 Usability measures

- Choice of measures. At least one measure should be provided for each of effectiveness, efficiency and satisfaction. There is no general rule on selection of the measures.

- Effectiveness. Goals of the user to the accuracy and completeness with which these goals can be achieved.
Efficiency. Measures related to the level of effectiveness achieved to the expenditure of resources.

Satisfaction. Measures related to the extent to which users are free from discomfort and their attitudes towards the use of the product.

3.3.1.4 Interpretation of measures

Generalization of the results should be done with care, especially when applying another context of use, which could have different types of users, tasks or environments.

Standard also provides examples, which are included in annexes and are of informative not prescriptive nature [1].

3.4 Usability engineering

Nielsen emphasizes that the usability engineering is a set of activities that in ideal situations take place throughout the lifecycle of the product and is not to be considered as a one-time activity to fix the user interface before the release of the product.

He proposes the following model in usability activities:

1. Know the user
2. Competitive analysis
3. Setting usability goals
4. Parallel design
5. Participatory design
6. Coordinated design of the total interface
7. Apply guidelines and heuristic analysis
8. Prototyping
9. Empirical testing
10. Iterative design
11. Collect feedback from field of use [17].

Usability also applies to the growth and expansion of a product and its portfolio, where products are released in many versions over a certain period of time. This broader view of usability justifies the need for allocation of usability engineering resources as early as possible. Collecting feedback from users and field of use has a much greater impact on future product releases as perceived by management.

Another author that addresses the usability engineering lifecycle and importance in understanding it as a set of activities is Mayhew. Her proposed model however is more focused on the process of designing good user interfaces for web pages and applets and includes also organizational and managerial strategies. Model proposes the following activities:

1. User profile
2. Contextual task analysis
3. Usability goal setting
4. Platform capabilities/constraints
5. General design principles
6. Workflow reengineering
7. Conceptual model design
8. Conceptual model mock ups
9. Iterative conceptual model testing
10. Screen design standards
11. Screen design standard prototyping
12. Iterative screen design standards testing
13. Style guide
14. Detailed user interface design
15. Iterative detailed user interface design testing
16. User feedback [18].

Even though there are distinctions between the two models and techniques associated with completing these activities, both put knowing your user and user feedback as an important part of the lifecycle.

### 3.5 Usability evaluation methods

Usability inspection has become more and more important as a method of evaluating user interfaces. Four basic ways of evaluation include automatically, empirically, formally and informally. The purpose of usability inspection is in finding usability problems in a design and overall usability of an entire system.

- **Heuristic evaluation.** It involves evaluators examining the interface and judging whether elements of it comply with the established usability principles.

- **Cognitive walkthroughs.** Use of detailed procedures to simulate the user’s problem solving process at each step through the dialog, to discover if the simulated users goals and memory can be assumed to lead to the next correct action.

- **Formal usability inspections.** Six-step procedure with defined roles in order to combine heuristic evaluation and a simplified for of cognitive walkthrough.

- **Pluralistic walkthroughs.** Meetings with users, developers and other stakeholders where they step through a scenario and discuss each element of the system.

- **Feature inspection.** Use case that is used to accomplish typical tasks, checks for long processes, cumbersome steps, not executable steps for users to try and steps that require extensive knowledge or experience in order to assess the proposed use case.

- **Consistency inspection.** Includes designers that represent multiple projects, inspect an interface, to establish whether the interface does the same things in the same way as their own design.

- **Standards inspection.** Includes compliance of an interface in relation to interface standard [19].
Most commonly used method is heuristic evaluation and Nielsen Norman group provided the ten usability heuristics for interface design which are still in use today. Author does however advise to use these principles more as a guideline, not specific usability guidelines 18.

Heuristic evaluation as one of the usability inspection methods has defined guidelines, which have been present and well established in the field by Nielsen. Heuristics include guidelines such as:

- Visibility of system status
- User control and freedom
- Error prevention
- Recovery from errors
- Flexibility and efficiency of use
- Aesthetics and minimalist design
- Recognition rather than recall
- Match between the system and the real world
- Help and documentation
- Consistency and standards.

Listed are the most general principles for interaction design [33]. The advantage of a heuristic evaluation is in the ability to provide a quick, inexpensive feedback to the designers; they can be used together with other usability testing methods. Downsides of such evaluation is in the expertise required for successful application, the use of a larger number of experts in order to aggregate the results and identification of more minor issues and few major ones 19.

These methods have also received some criticism within the actual practice of system development. Wixon argues these scientific methods with the fact that, in the real world of product development, the schedule and the resource issues often dominate all discussions of design and benefits, due to the nature of these issues and because the product goals are not clear enough or perceived to be more speculative. Usability work within the industry the focus is on factors of success – how effective is the introduced method in improving products. He emphasizes the case study approach is the most effective way to produce a body of knowledge for applied usability. Arguments are illustrated with two case studies which conclude that literature evaluating usability methods is flawed in terms of industry and proposes a broad-based range case study approach in order to provide relevant outcomes both for practice and business [20].

18 http://www.nngroup.com/articles/ten-usability-heuristics/ (Visited April 2014)
3.6 The distance problem

The focus if this thesis is to investigate the distance problem. Methods of usability within different research communities address the mass market solutions and a wide range of users. As such are in need of adjustment to capture the context of use within specific systems (M-Commerce as an example) and expert users that are dispersed over the world.

Personas are widely used in user experience research and usability and they are defined as a single, fictitious person who represents the needs and requirements of many people\textsuperscript{20}. They are built on interviews with actual users of systems. Their use can be helpful to the design, but they can be more powerful if used to complement a full range of qualitative and quantitative methods and have the possibility to amplify the effectiveness of those [34].

Focus groups are used as an informal technique to assess the user needs and feelings before and after the interface has been in use. In a focus group, users are encouraged to discuss new concepts and usability issues discovered over a longer period of time [36]. Focus groups can also be subjected to remote testing, with the use of video conferencing and other forms of electronic networks, which is an inexpensive way of conducting usability testing.

As stated in the introduction, EWP is a newly developed system that has yet to emerge onto the market. While there are currently deployments ongoing at customer environments in Africa, a much bigger roll-out is expected over the coming years. Therefore the aspect of remote usability testing is to be looked at also.

In remote usability testing user research can be conducted with participants in their natural environment with the use of modern technology like screen sharing or online remote usability services. This type of testing is to be considered in conditions that include:

- Timelines preventing in-person testing due to scheduling issues
- Participants are geographically dispersed
- Use of particular work equipment due to software or security requirements
- Accessibility issues, which require that the users use their own software or equipment

Even though this method is flexible in applying, benefits and challenges of using such methodology should be considered as well.

Benefits:

- No need for a lab environment and the effect of a lab environment on participants
- Intended for diverse groups of participants
- Less expensive than traditional in-person lab testing
- In the case of un-moderated testing test day can be extended, allowing test to be subjected to a larger group of participants

\textsuperscript{20} http://thecontentwrangler.com/2011/08/23/personas-in-user-experience/ (Visited May 2014)
Opportunity to conduct the test with a larger group of people than a lab environment accommodates

Challenges:

- Security could be compromised
- Restricted or no view of the participants body language
- Technical difficulties (technology related)\(^\text{21}\).

The remote usability testing method is also used in order to get customer insights or when test participants are hard to find or located in different locations\(^\text{22}\).

After a literature review that addresses different methodologies in usability testing, a gap in the knowledge is detected. While different case studies show different applications of methodology and approaches in usability testing, there is a lack of knowledge in the area of M-Commerce systems. Systematic literature overview serves as input to the extension of usability testing method that considers the development and globally dispersed customers.

\[^{21}\text{http://www.usability.gov/how-to-and-tools/methods/remote-testing.html}\ (Visited May 2014)\]

\[^{22}\text{http://www.nngroup.com/articles/remote-usability-tests/}\ (Visited May 2014)\]
4 M-COMMERCE AND ERICSSON´S SOLUTION

The term M-Commerce has become widely used in the present time with the phrase introduced over 15 years ago on a global mobile commerce forum by Kevin Duffey. His suggested preliminary definition of M-Commerce was “the delivery of electronic commerce capabilities directly into the consumer’s hand, anywhere via wireless technology”.

4.1 M-Commerce

To get a better understanding of what M-Commerce is, some definitions need to be looked at.

According to Tsalgatidou, M-Commerce refers to e-commerce based activities depending only or partially on mobile e-commerce transactions. Under the term mobile e-commerce transaction, any type of transaction of an economic value that is conducted through a mobile device that uses wireless telecommunications network for communication with the e-commerce [25].

Another definition characterizes M-Commerce as any direct or indirect transaction conducted and facilitated through a wireless telecommunication network. M-commerce technologies have the potential to improve consumer’s lives, create new business and make existing ones more profitable [26].

Hanebeck defines M-Commerce as a delivery of electronic commerce capabilities directly to consumers via wireless technology enabling a retail outlet to be put into the consumers’ hand. Businesses reach the consumers directly and independently of their location with this type of commerce [27].

M-Commerce solutions consist of a robust ecosystem which enables:

- A transparent transaction process
- Covering all critical stakeholders
- Quick feedback, clearing, settlement and risk management process
- Regulatory framework
- Strong data security framework.

23 [http://cryptome.org/jya/glomob.htm](http://cryptome.org/jya/glomob.htm) (Visited May 2014)
Key stakeholders in this solution are subscribers, Mobile Network Operators (MNO), merchants, retailers, banks, micro finance institutions, service industries and government institutions. Their relationships in the M-Commerce are illustrated in the picture below:

Figure 3: M-Commerce ecosystem\(^{24}\) (source: Solution Description, Ericsson)

### 4.2 Ericsson Wallet Platform

As mentioned previously, Ericsson is with a solution for M-Commerce emerging onto the market and the system is developed in Karlskrona. Ericsson Wallet Platform (EWP) is a solution that addresses a broad range of different stakeholders connected with mobile money services. EWP features include:

- Open platform - design services for multiple consumer segments.
- Enhanced user experience – with the deployment of better user interfaces that can lead to a more active wallet.
- Easy integration - seamlessly merge with existing technology.
- Complete security – connect securely to the ecosystem around the MNO
- Built-in compliance - adheres to all rules and regulations.

Solution has two possibilities of deployment; as a solution for domestic mobile financial services but can also be connected to international remittance providers. It can function as a self-contained platform, but can also operate as the Mobile Money module in the Ericsson Converged Wallet (ECW) solution which extends the Ericsson Charging system.

Key capabilities of the solution are:

- One account view
- Cross offering promotions
- One configuration tool for products and promotions, and transaction tariff rules.

Users that can access the system have different roles and they perform different types of operations within the system in order to enable mobile money services.

Figure 4: Stakeholders in EWP (source: Solution Description, Ericsson)

Users are accessing the system using various types of interfaces and as mentioned before, the object of usability testing is the graphical user interface that supports the work of customer care agents, financial controllers and compliance officers.
4.3 The interface – Customer Care GUI

Customer service support is executed via a web based interface called the Customer Care GUI, giving customer care representatives a tool to manage inquiries from new and existing users, merchant channels, registered agents and partners. The interface is profile driven and the users accessing the system can only perform authorized tasks assigned to the profile. Profiles include access to the following tasks:

- Account holder management
- Sanctions
- Financial controller
- Chanel distribution
- Approval

Profiles employ different types of users with different levels of expertise and knowledge and are considered to be the expert users of the system.

Figure 5: Customer Care GUI - Account holder view*

*Due to the sensitive nature of information (account holder data) presented in the interface, the image is blurred.
4.4 Usability in the interface

The interfaces to access the solution have undergone an exploratory testing method. Method is considered to be ad hoc and can be explained as “any testing to the extent that the tester actively controls the design of the tests and uses information gained while testing to design new and better tests” [28].

Although the exploratory method is commonly used in software development, usability testing has not yet been performed on the interface of EWP solution. The aim with this thesis project is in evaluating established usability testing methods and to possibly extend the method with addressing the distance problem. Method, if proven effective will be a part of usability framework used in Ericsson and also applied at customer sites in the future.

4.5 Case study

Previous chapters gave an overview of the system and different types of users that are considered to be expert users and this usability test is intended to measure the context of use of the expert user group by performing four tasks used on a daily basis. As previously mentioned, the object of this investigation is the Customer Care GUI, which supports the work of customer care personnel at MNO’s offices. The main purpose of this usability test was in measuring the context of use in the Customer Care GUI with the test participants from the SA-team acting as proxies for the expert users of the interface and to detect possible usability issues. Using this technique, a scenario with realistic situation (task to be completed uses only one interface, to observe if it is successfully completed with a level of satisfaction) was created and test participants performed tasks while I observed them and took notes (for more information about the test cases please refer to appendix A). Testing techniques such as instructions and a post-test questionnaire is used to collect feedback from the test participants. Methods of evaluation that included the TAP protocol and remote testing were used to gather data during the execution of the test.

4.6 Usability test design

After a consultation with the user experience department at Ericsson and the knowledge obtained from the literature study, a usability test was designed. It combined a set of tasks that are measured (completion times, number of errors) with the application of the TAP protocol and cooperative evaluation [30] as advised by external supervisor. Interviews were conducted with the test participants and a questionnaire was designed in order to steer the interviewing process. The elements used in this usability test and their theoretical foundations are presented in this chapter.

Questionnaire is a set of questions for gathering information from individuals. They can be administered in different ways and they are commonly used when resources are limited and data is needed from people, in order to gather specific data like knowledge, attitudes and behavior or when the privacy of the participants is a concern (in the case of sensitive data collection – participants responses can be anonymous) [25]. It serves as an inexpensive method to collect different types of data in a fast and efficient way. However it is of importance to design the questionnaire well in order to achieve clarity and questions have the same meaning to different types of respondents.

Semi-structured interview combines some structured questions with some unstructured exploration. The general goal of this type of interview is in gathering systematic information about a set of central topics, while also allowing some exploration when new issues arise. They are used when there is some knowledge about the topic of investigation, but further details are still required. Main advantages include:

- Possibility to uncover previously unknown issues
- Addressing complex topics through clarification and iterations
- Provide some flexibility and allows some broad comparisons across interviews
- Require less training time due to the fact that the interviewer has a specific set of questions available.

While the advantages are great, the disadvantages of this type of interview should also be considered. As weaknesses of a semi-structured interview the following raises a concern:

- Some training and experience is required, so that the interviewer does not put words into the participant’s mouth
- Interviewers can give clues that guide the participant in a particular answer.
- The mixture of quantitative and qualitative data that is the result of the interview can be time consuming to analyze and might prove to be hard to generalize.

Despite the weaknesses of a semi-structured interview, this type of interviews are recommended to be used in situations where understanding of user goals, gathering information about tasks, task flow and work artifacts, gathering data on topics when the interviewer is certain to some degree that the relevant issues have been identified, but still provide respondents with the possibility to address new issues that are important to them through open-ended questions [29].

A questionnaire addressing the usability of the investigated interface was formulated in order to assist in the semi-structured interview. All test participants were asked the same questions and their opinion was gathered as well.

Think aloud protocol is considered to be the number one usability tool, since it employs simple usability tests, where the users think out loud, they are inexpensive robust and easy to learn. However the use of the TAP protocol has some risks involved and it does not resolve all the problems. Nielsen has defined the method as a test that has the test participant use the system while continuously thinking out loud. With verbalization of participant’s thoughts the test users enable the tester to understand how they view the computer system and this serves as a way to identifying users major misconceptions [30].

The TAP is a very direct method to gain insight in the knowledge and methods of human problem-solving. The method can be used by psychologists and other social scientists that

---

26 http://www.nngroup.com/articles/thinking-aloud-the-1-usability-tool/ (Visited May 2014)
want to know more about the cognitive processes; and is also is an important method for usability engineers whose goal is to build a system on the basis of human expertise [31].

The protocol was applied thought the entire usability test, where the participants were reminded to provide critical insight of the interface functionality and to also to think from the expert group perspective.

While the method has advantages like simplicity, little expertise to perform with the capability to provide a useful insight to problems with an interface, there are a number of problems that are associated with it. The information is subjective – being observed and having to describe what you are doing can also affect the way you complete a task. A variation of the TAP method is “Cooperative evaluation” which extends the method with encouraging the user to see himself as a collaborator in the evaluation as opposed to just a subject. The participant is encouraged to actively criticize the system rather than to simply perform the tasks given and think out loud. Problems associated with both techniques are in generation of a large volume of data, which can be time consuming to analyze²⁷.

To summarize the described components that were used, the usability test consisted of the following steps:

- test preparation
- test participant selection
- task definition – the test cases test participants had to perform
- test execution.

Steps are defined and described in detail in the following subchapters.

### 4.6.1 Test preparation

The usability test was conducted by evaluator Gabriela N. The test consisted of a number of test cases which the participants had to perform (described in sub-chapter 4.6.3). The test began with a short introduction and presentation of this thesis project and continued in a brief description of tasks, techniques used to collect data and the aim of this usability test. This was presented to each participant. After the test participants completed the tasks and the times to complete the tasks were measured and recorded, the interview process began. It was a semi-structured interview based on the questionnaire (see appendix A for details) and after the interview was finished, the usability test was concluded. Gathered data was used for analysis which is presented in chapter 5.

Figure 6: Usability test overview

### 4.6.2 Test participants

As a first step in usability engineering lifecycle model proposed by Nielsen, knowing the user is of critical importance [17]. It consists of studying the intended users and the use of the product. He suggests that the developers should at some point visit the customer’s premises in order to get an understanding of how the system will be used. Individual user’s characteristics and various tasks are two factors that have the largest impact on usability. When users are considered, the concept of a user should be extended to include everybody whose work is affected by the system, even if they never see an interface.

In most cases it is difficult for developers to get access to intended users of the developed system in order to fulfill the basic usability guideline – know the user. Grudin examines the problems in such access which includes:

- Protection of the development team by the company in order to prevent possible customer bypasses – direct calls to developers, which results in sidetracking them from their main development job

- Sales representatives and their reluctance to allow anybody else from the company to talk to the customers in order not to potentially offend the customer or create dissatisfaction.

- Short amount of time where the users are available [32].

SA-team at Ericsson is required to visit customer sites and gather information; their output is in providing documentation such as solution description, user acceptance test cases and
various reports. Participants in this usability test are listed in the table below (due to security reasons according to NDA document of Ericsson, their full names cannot be mentioned):

<table>
<thead>
<tr>
<th>Number</th>
<th>Name</th>
<th>Experience in the SA job role in years</th>
<th>Sex</th>
<th>Age category</th>
<th>Date of the usability test</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Glendale G</td>
<td>less than 3</td>
<td>Female</td>
<td>35-40</td>
<td>19.05.2014</td>
</tr>
<tr>
<td>2</td>
<td>Stefan W</td>
<td>over 10</td>
<td>Male</td>
<td>40-45</td>
<td>19.05.2014</td>
</tr>
<tr>
<td>3</td>
<td>Åke J</td>
<td>over 10</td>
<td>Male</td>
<td>40-45</td>
<td>20.05.2014</td>
</tr>
<tr>
<td>4</td>
<td>Srdjan K</td>
<td>less than 3</td>
<td>Male</td>
<td>40-45</td>
<td>23.05.2014</td>
</tr>
<tr>
<td>5</td>
<td>Pavan B</td>
<td>over 5</td>
<td>Male</td>
<td>30-35</td>
<td>26.05.2014</td>
</tr>
</tbody>
</table>

Table 1: Test participants

In knowing the user, individual user characteristics must be studied. In some cases this is easily achieved since it is possible to identify users as concrete individuals (when a system is used in a specific department in a specific company). For some systems, users can be widely scattered and it is possible to visit only a few, representative customers [11].

Test participants chosen to perform the usability test for this thesis project were selected among the SA-team, as one of their job duties is in collecting requirements for system adjustments and configuration, requirements that expand the system in order to support specific tasks discovered as potentially beneficial in EWP solution. Their work and customer interaction provides insights to actual use of the EWP solution and they therefore serve as proxies for the real users. The solution architect job role is considered to be one of most sought out roles in software development process. The main purpose of the role is in conversion of the requirements into an architecture and design that will become a blueprint of the created solution.\(^\text{28}\)

4.6.3 Task definition

Interviews and discussions with colleagues in regards to their work on customer sites and communication with the expert group and work processes helped me to select 4 tasks to be used in the usability test and the SA-team reviewed the relevance of the selection made.

Tasks can be characterized as the activities undertaken to achieve a goal. Description of the task should not be limited solely in terms of functions or features provided by a system. Detailed description of the activities and steps required to perform and complete the given task must also be documented [4]. In this process the SA-team plays an important role, as they participate in workshops with the expert users and other stakeholders in order to document work processes and goals that are to be achieved with these tasks.

Task analysis serves as essential early input to system design. Overall goals of the intended users should be studied as well as current executions of the tasks, their information needs and how deviations normal circumstances or emergencies are handled. Common outcome of the task analysis is a list of everything that the intended users want to accomplish with the use of the system [11].

For the purpose of this usability test four tasks performed on a daily basis by the expert group at MNO’s offices were selected:

- Transfer of funds. An amount of 10 EUR to be transferred between account holder A and account holder B.
- Edit account holder information. User is to search for an account holder and edit the account holder information – update of the street number in the address field.
- View transaction history. User is to search for an account holder and for the specific account holder select view historical data and view vouchers.
- 4 eye principle. The four eye principle is a process that prevents users from performing certain operations in a single step. An approval is required by another user, for example a supervisor.

Test tasks and the steps necessary to execute them are presented in appendix B.

4.6.4 Test execution

The usability test was conducted at Ericsson, Karlskrona. Participants received individual invitations to the usability test session via Outlook tool and team rooms were booked for the session in order to prevent possible interruptions and distractions. The printed out papers of test cases to be executed were supplied to the participants. Participants used the evaluator’s laptop, while one used his own, in order to access the environment containing the latest EWP installation with the interfaces. The environment was accessed via Internet Explorer, which is the only supported web browser in Ericsson. A post usability questionnaire was printed out and provided to the users in order to steer the interview process and gather feedback, while the participant performing the usability test remotely received his questionnaire via e-mail.
5 TEST RESULTS AND ANALYSIS

In the usability test, each test participant performed four tasks according to the test case specification (see appendix B for detailed description). Participants were asked to “think aloud” and provide comments and recommendations regarding the interface during execution of the tasks. Upon completion, the interview process began. As mentioned in the previous chapter, the interview was steered by a questionnaire, in order for all participants to address the same questions regarding efficiency, effectiveness and satisfaction. Open discussion regarding the testing method, data collection and experiences from customer sites was present at all time.

The quantitative (numeric values) and the qualitative data (non-numeric) collected was analyzed and the results are presented for separate test cases, to combine both types of data and provide conclusions in regards to the usability in the interface.

For each test case measurements of time completion, number of errors, time to support the participant and test case categorization completed or not completed were recorded. They are presented in the table below:

<table>
<thead>
<tr>
<th>PARTICIPANT ID</th>
<th>TOTAL TIME TO COMPLETE</th>
<th>TIME TO SUPPORT</th>
<th>TIME WITHOUT SUPPORT</th>
<th>NUMBER OF ERRORS</th>
<th>COMPLETED/NOT COMPLETED</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>14:36</td>
<td>01:08</td>
<td>13:28</td>
<td>1</td>
<td>Completed</td>
</tr>
<tr>
<td>2</td>
<td>16:24</td>
<td>02:00</td>
<td>14:24</td>
<td>2</td>
<td>Completed</td>
</tr>
<tr>
<td>3</td>
<td>08:32</td>
<td>00:00</td>
<td>08:32</td>
<td>0</td>
<td>Completed</td>
</tr>
<tr>
<td>4</td>
<td>13:05</td>
<td>01:20</td>
<td>11:45</td>
<td>1</td>
<td>Completed</td>
</tr>
<tr>
<td>5</td>
<td>18:50</td>
<td>03:50</td>
<td>15:00</td>
<td>3</td>
<td>Completed</td>
</tr>
</tbody>
</table>

Table 2: Usability test result

All five participants (as mentioned before, according to Nielsen, the optimal number of testers is five\textsuperscript{29}) completed the test in less than 20 minutes, participant nr. 3 completed the test in the shortest amount of time – in less than 10 minutes (due to the fact she had an active role in development). Number of errors per participant did not exceed 3 and time spent to support the participants was between 1 and 4 minutes.

\textsuperscript{29} http://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users/ Visited April 2014
The total time each participant spent in task execution is presented in the form of a graph, where the blue color presents the total time to complete and red color the time to support in order to complete the test cases. As seen from the graph, participant 3 performed the given tasks with no errors and with the shortest amount of time.

![Time to complete (total)](image)

Figure 7: Time to complete the task

### 5.1 Usability test results – UTC01 – Transfer of funds

The test case covers a daily activity performed by the expert user of the interface. The results gathered are presented in the table:

<table>
<thead>
<tr>
<th>PARTICIPANT ID</th>
<th>TEST CASE 01 - TIME TO COMPLETE</th>
<th>TIME TO SUPPORT</th>
<th>TIME WITHOUT SUPPORT</th>
<th>NUMBER OF ERRORS</th>
<th>COMPLETED/NOT COMPLETED</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>03:42</td>
<td>01:08</td>
<td>02:34</td>
<td>1</td>
<td>Completed</td>
</tr>
<tr>
<td>2</td>
<td>05:53</td>
<td>01:01</td>
<td>04:52</td>
<td>2</td>
<td>Completed</td>
</tr>
<tr>
<td>3</td>
<td>02:29</td>
<td>00:00</td>
<td>02:29</td>
<td>0</td>
<td>Completed</td>
</tr>
<tr>
<td>4</td>
<td>03:40</td>
<td>00:50</td>
<td>02:50</td>
<td>1</td>
<td>Completed</td>
</tr>
<tr>
<td>5</td>
<td>06:52</td>
<td>01:50</td>
<td>05:02</td>
<td>3</td>
<td>Completed</td>
</tr>
</tbody>
</table>

Table 3: UTC01 test results

The task proved to be challenging, since four participants needed support, due to the fact the interface was not self-explanatory as to where to find the transfer option. After support was provided the test case was successfully completed. Participants expressed the need for revising the terms used for actions and inconsistency with input fields was noted as well. As an alternative, a suggestion of the transfer action to be added to the account holder information was made.
5.2 Usability test results – UTC02 – Edit account holder information

The test case consisted of several steps and the participants were all able to complete the given task with no errors or help required.

<table>
<thead>
<tr>
<th>PARTICIPANT ID</th>
<th>TEST CASE 02 - TIME TO COMPLETE</th>
<th>TIME TO SUPPORT</th>
<th>TIME WITHOUT SUPPORT</th>
<th>NUMBER OF ERRORS</th>
<th>COMPLETED/NOT COMPLETED</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>01:02</td>
<td>00:00</td>
<td>01:02</td>
<td>0</td>
<td>Completed</td>
</tr>
<tr>
<td>2</td>
<td>01:39</td>
<td>00:00</td>
<td>01:39</td>
<td>0</td>
<td>Completed</td>
</tr>
<tr>
<td>3</td>
<td>02:12</td>
<td>00:00</td>
<td>02:12</td>
<td>0</td>
<td>Completed</td>
</tr>
<tr>
<td>4</td>
<td>06:52</td>
<td>00:00</td>
<td>06:52</td>
<td>0</td>
<td>Completed</td>
</tr>
</tbody>
</table>

Table 4: UTC02 test results

Despite the fact the test case was completed, there was an improvement suggestion, which was made based on the experiences with the expert users and their expectations of this particular task. A suggestion of a more configurable view was also recorded (expert user can control the tasks and in which order they are displayed).

5.3 Usability test results – UTC03 – View transaction history

As a part of account holder management, a view of transaction and voucher history is presented and the objective of this test case. Usability issues were detected with not all participants knowing where to start and the interface not being self-explanatory was highlighted again.

<table>
<thead>
<tr>
<th>PARTICIPANT ID</th>
<th>TEST CASE 03 - TIME TO COMPLETE</th>
<th>TIME TO SUPPORT</th>
<th>TIME WITHOUT SUPPORT</th>
<th>NUMBER OF ERRORS</th>
<th>COMPLETED/NOT COMPLETED</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>05:01</td>
<td>00:00</td>
<td>05:01</td>
<td>0</td>
<td>Completed</td>
</tr>
<tr>
<td>2</td>
<td>03:30</td>
<td>00:59</td>
<td>02:31</td>
<td>1</td>
<td>Completed</td>
</tr>
<tr>
<td>3</td>
<td>02:52</td>
<td>00:00</td>
<td>02:52</td>
<td>0</td>
<td>Completed</td>
</tr>
<tr>
<td>4</td>
<td>03:52</td>
<td>00:50</td>
<td>03:02</td>
<td>0</td>
<td>Completed</td>
</tr>
<tr>
<td>5</td>
<td>05:57</td>
<td>01:50</td>
<td>04:07</td>
<td>3</td>
<td>Completed</td>
</tr>
</tbody>
</table>

Table 5: UTC03 test results

The display of transaction history was confusing and hard to understand, which indicates the particular task should be revised and a better solution provided.
5.4 Usability test results – UTC04 – Four-eye-principle

The enforcement of the four-eye-principle is a very common part of tasks within the financial industry, due to the possibility of errors or missuses. It consists of special permissions in order for some actions to be approved by a supervisor.

<table>
<thead>
<tr>
<th>PARTICIPANT ID</th>
<th>TEST CASE 04 - TIME TO COMPLETE</th>
<th>TIME TO SUPPORT</th>
<th>TIME WITHOUT SUPPORT</th>
<th>NUMBER OF ERRORS</th>
<th>COMPLETED/NOT COMPLETED</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>04:51</td>
<td>01:08</td>
<td>03:43</td>
<td>0</td>
<td>Completed</td>
</tr>
<tr>
<td>2</td>
<td>05:07</td>
<td>00:59</td>
<td>04:08</td>
<td>1</td>
<td>Completed</td>
</tr>
<tr>
<td>3</td>
<td>02:52</td>
<td>00:00</td>
<td>02:52</td>
<td>0</td>
<td>Completed</td>
</tr>
<tr>
<td>4</td>
<td>03:52</td>
<td>00:50</td>
<td>03:02</td>
<td>0</td>
<td>Completed</td>
</tr>
<tr>
<td>5</td>
<td>05:57</td>
<td>01:50</td>
<td>04:07</td>
<td>3</td>
<td>Completed</td>
</tr>
</tbody>
</table>

Table 6: UTC04 test results

The test case was completed with some errors and comments were gathered. This task proved to be the most confusing one to the users and a great amount of improvement possibilities was detected, such as approval history, information of pending approval tasks being displayed for all users, and more logical positioning of the approval action.

5.5 Usability test observation

All of the participants were encouraged through the usability test to express their opinions and the TAP [31] method with its extended variant of the cooperative evaluation served as a powerful tool into the insights and problems the SA-team faces when interacting with the expert users. In the case of the remote test, it was not possible to gather and document all observations of the participant. The Microsoft conference tool Lync was used as a standard within Ericsson and it does not enable a full view of the room and gestures of the user were not seen to the extent as in test conduction in Karlskrona.

As mentioned in the introduction chapters, no usability testing was conducted on the interfaces and the test participants expressed enthusiasm with this thesis project and were therefore eager to help.

The interview that followed the usability test was beneficial to all involved; including myself, as discussions on various usability issues and how to avoid them in the future releases of the EWP solution, gave me an insight into the expert group of users and their work processes.

Based on the usability test conduction and observations gathered a list of recommendations for the development team was created. The entire test was also beneficial to the test users, since they were acting as proxies for the expert users at customer sites and they were able to share difficulties that they themselves observed while working with the expert group.

It is of vital importance to employ expert users that will use the developed system, as they are the ones that use it as a part of their daily work. The proposed method therefore suggests the use of the SA-team, which acts as a proxy to the expert user located in Africa.
6 DISCUSSION

6.1 Proposed usability method evaluation – internal

Usability testing and user experience within Ericsson’s department in Karlskrona consists of following guidelines and conduction of tests that include:

- usability test setup
- preparation of the test
- conduction of the test based on best practices
- analysis and test reporting.

The data gathered addresses various aspects of the system and is conducted within the development department. In the cases where other departments are involved in the process the SUS questionnaire is used as well, since the method is an industrial standard and most used in measuring perceptions of usability.\(^{30}\)

The context of use based usability testing is not a common and established practice within the department and the initiative in the form of this thesis project was welcomed. The results and recommendations provided are under review and will be examined with the potential re-use of the method and its possible generalizations for use in the development department.

As a part of EWP solution, training course for the actual users of the system is also available. Training is commonly conducted at customer sites, but due to the personnel in need of training coming from various countries, training course was held in Karlskrona. Having the actual users of the system present in the company, another set of usability testing was conducted. Usability test conducted with the SA-team was repeated with the same conditions and environment, but with the actual users. Second test served as an assessment of the proposed method and its results are presented in this subchapter.

The interface used for this usability test is used by customer care personnel at a mobile network operator (MNO) office and enables employees to perform different types of account and user management. Interface is used daily as a part of duties and tasks of an employee of MNO. Participants in the second iteration of the test are listed in the table below (due to security reasons according to NDA document of Ericsson, their full names cannot be mentioned):

Table 7: Test participants – actual users

<table>
<thead>
<tr>
<th>Number</th>
<th>Name</th>
<th>Experience in the job role in years</th>
<th>Sex</th>
<th>Age category</th>
<th>Date of the usability test</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Olivier D.</td>
<td>Over 15</td>
<td>Male</td>
<td>45-50</td>
<td>23.06.2014</td>
</tr>
<tr>
<td>2</td>
<td>Lanre A.</td>
<td>Less than 5</td>
<td>Male</td>
<td>30-35</td>
<td>19.06.2014</td>
</tr>
<tr>
<td>3</td>
<td>Thomas B.</td>
<td>Over 15</td>
<td>Male</td>
<td>40-45</td>
<td>19.06.2014</td>
</tr>
<tr>
<td>4</td>
<td>Laura M.</td>
<td>Over 15</td>
<td>Female</td>
<td>40-45</td>
<td>18.06.2014</td>
</tr>
</tbody>
</table>

For each test case (as in previous iteration) measurements of time completion, number of errors, time to support the participant and test case categorization completed or not completed were recorded. They are presented in the table below:

Table 8: Usability test result – actual users

<table>
<thead>
<tr>
<th>PARTICIPANT ID</th>
<th>TOTAL TIME TO COMPLETE</th>
<th>TIME TO SUPPORT</th>
<th>TIME WITHOUT SUPPORT</th>
<th>NUMBER OF ERRORS</th>
<th>COMPLETED/NOT COMPLETED</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>15:42</td>
<td>02:04</td>
<td>13:38</td>
<td>4</td>
<td>Completed</td>
</tr>
<tr>
<td>2</td>
<td>13:19</td>
<td>01:05</td>
<td>12:14</td>
<td>2</td>
<td>Completed</td>
</tr>
<tr>
<td>3</td>
<td>12:32</td>
<td>02:00</td>
<td>10:32</td>
<td>0</td>
<td>Completed</td>
</tr>
<tr>
<td>4</td>
<td>11:25</td>
<td>01:20</td>
<td>10:05</td>
<td>1</td>
<td>Completed</td>
</tr>
</tbody>
</table>

Four participants (as mentioned before, according to Nielsen, the optimal number of testers is five\(^{31}\)) completed the test in less than 20 minutes, participant nr. 4 completed the test in the shortest amount of time – in less than 12 minutes. Number of errors per participant did not exceed 4 and time spent to support the participants was between 1 and 3 minutes.

Graphical representation of results is given below, where the blue color presents the total time to complete and red color the time to support in order to complete the test cases. As seen from the graph, all participants needed some support in test case execution.

Figure 8: Time to complete the tasks - actual users

6.1.1 Usability test results – UTC01 – Transfer of funds

The test case covers a daily activity performed by the expert user of the interface. The results gathered are presented in the table:

<table>
<thead>
<tr>
<th>PARTICIPANT ID</th>
<th>TEST CASE 01 - TIME TO COMPLETE</th>
<th>TIME TO SUPPORT</th>
<th>TIME WITHOUT SUPPORT</th>
<th>NUMBER OF ERRORS</th>
<th>COMPLETED/NOT COMPLETED</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>03:34</td>
<td>01:00</td>
<td>02:34</td>
<td>3</td>
<td>Completed</td>
</tr>
<tr>
<td>2</td>
<td>04:43</td>
<td>00:30</td>
<td>04:13</td>
<td>1</td>
<td>Completed</td>
</tr>
<tr>
<td>3</td>
<td>03:29</td>
<td>01:00</td>
<td>02:29</td>
<td>0</td>
<td>Completed</td>
</tr>
<tr>
<td>4</td>
<td>03:50</td>
<td>00:50</td>
<td>03:00</td>
<td>1</td>
<td>Completed</td>
</tr>
</tbody>
</table>

Table 9: UTC01 test results – actual users

Similar results as with the SA-team were obtained with the actual users. Interface was not self-explanatory and users did not know where to find the transfer option. After provided support, the test case was successfully completed. Revision of the terms was noted as an improvement and one of the participants also pointed out the need to avoid scrolling in order to complete the action.

6.1.2 Usability test results – UTC02 – Edit account holder information

The test case consisted of several steps and the participants were all able to complete the given task.

<table>
<thead>
<tr>
<th>PARTICIPANT ID</th>
<th>TEST CASE 02 - TIME TO COMPLETE</th>
<th>TIME TO SUPPORT</th>
<th>TIME WITHOUT SUPPORT</th>
<th>NUMBER OF ERRORS</th>
<th>COMPLETED/NOT COMPLETED</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>03:05</td>
<td>01:00</td>
<td>02:05</td>
<td>1</td>
<td>Completed</td>
</tr>
<tr>
<td>2</td>
<td>02:33</td>
<td>00:30</td>
<td>01:54</td>
<td>1</td>
<td>Completed</td>
</tr>
<tr>
<td>3</td>
<td>02:45</td>
<td>00:00</td>
<td>02:45</td>
<td>0</td>
<td>Completed</td>
</tr>
<tr>
<td>4</td>
<td>02:20</td>
<td>00:00</td>
<td>02:20</td>
<td>0</td>
<td>Completed</td>
</tr>
</tbody>
</table>

Table 10: UTC02 test results – actual users

Test case was completed with minor support provided to the first two users, improvement suggestions was also noted – search button should be present, going back button to be added and system message of successful edit should be displayed as well.
6.1.3 Usability test results – UTC03 – View transaction history

As a part of account holder management, a view of transaction and voucher history is presented and the objective of this test case. As with the SA-team, some usability issues were detected and noted.

<table>
<thead>
<tr>
<th>PARTICIPANT ID</th>
<th>TEST CASE 03 - TIME TO COMPLETE</th>
<th>TIME TO SUPPORT</th>
<th>TIME WITHOUT SUPPORT</th>
<th>NUMBER OF ERRORS</th>
<th>COMPLETED/NOT COMPLETED</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>05:19</td>
<td>00:00</td>
<td>05:19</td>
<td>0</td>
<td>Completed</td>
</tr>
<tr>
<td>2</td>
<td>01:56</td>
<td>00:00</td>
<td>01:56</td>
<td>0</td>
<td>Completed</td>
</tr>
<tr>
<td>3</td>
<td>03:53</td>
<td>01:00</td>
<td>02:53</td>
<td>0</td>
<td>Completed</td>
</tr>
<tr>
<td>4</td>
<td>02:30</td>
<td>00:00</td>
<td>02:30</td>
<td>0</td>
<td>Completed</td>
</tr>
</tbody>
</table>

Table 11: UTC03 test results – actual users

The display of transaction history was confusing and hard to understand, suggestion of explanation of the colors used in the table was recorded. One participant also expressed the need to make the data in the transaction history exportable as a pdf file format.

6.1.4 Usability test results – UTC04 – Four-eye-principle

The enforcement of the four-eye-principle is a very common part of tasks within the financial industry, due to the possibility of errors or missuses. It consists of special permissions in order for some actions to be approved by a supervisor.

<table>
<thead>
<tr>
<th>PARTICIPANT ID</th>
<th>TEST CASE 04 - TIME TO COMPLETE</th>
<th>TIME TO SUPPORT</th>
<th>TIME WITHOUT SUPPORT</th>
<th>NUMBER OF ERRORS</th>
<th>COMPLETED/NOT COMPLETED</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>03:44</td>
<td>00:04</td>
<td>03:40</td>
<td>0</td>
<td>Completed</td>
</tr>
<tr>
<td>2</td>
<td>04:07</td>
<td>00:35</td>
<td>03:32</td>
<td>0</td>
<td>Completed</td>
</tr>
<tr>
<td>3</td>
<td>02:25</td>
<td>00:00</td>
<td>02:25</td>
<td>0</td>
<td>Completed</td>
</tr>
<tr>
<td>4</td>
<td>01:45</td>
<td>00:30</td>
<td>01:15</td>
<td>0</td>
<td>Completed</td>
</tr>
</tbody>
</table>

Table 12: UTC04 test results – actual users

The test case was completed with no errors and comments were gathered. This task proved to be the most confusing one to the users and their thoughts on improvement were similar to the thoughts of the SA-team.

6.1.5 Usability test summary

The time to complete the first test case was approximately the same for the SA-team members and the actual users. The SA-team as well as the actual users expressed the need for revising the terms used for actions, and an inconsistency with input fields was noted as well.

In the second test case, some improvement suggestions were noted from both groups of users (SA-team and actual users): a search button should be present, a going back button should be added and a system message of successful edit should be displayed.

In the third test case, both groups thought that the display of transaction history was confusing and hard to understand, and suggestions of explanation of the colors used in the
table were recorded. One actual user also expressed the need to make the data in the transaction history exportable as a PDF file.

The fourth test case was the most confusing one for the SA-team as well as for the actual users and many improvement possibilities were detected, such as approval history, information of pending approval tasks being displayed for all users, and more logical positioning of the approval action.

The second iteration of the usability test with the actual users served as an evaluation of the proposed initial method, which included the SA-team acting as proxies. As is evident from the results (see tables 2-6, and tables 8-12), there was no major deviation in completion times or number of errors. While some support was provided to the test participants, all were able to complete the tasks and express their thoughts and opinions. Both groups completed the entire test in less than 20 minutes. The main purpose with the internal validation was achieved, as the actual users gave similar test results and similar comments regarding the interface (see summary above).

This study included five proxy users and four actual users. It is, of course, hard to say for sure how the results should be with a different set of users. However, five users is (as previously stated) the number recommended by Nielsen, and since the results obtained from the proxy users and the actual users were very similar, it is reasonable to assume that my results are reliable.

As stated in the conclusion of the case study conducted in the health sector in Norway, the ISO standard proves to be powerful guideline in the context of use testing [6], which was considered in this thesis work and has lead me to the same reasoning.

### 6.2 Proposed usability method evaluation – external

Usability activities in industrial organizations remain a neglected process and are as such also often sacrificed over functionality, time constraints and costs. However, products with poor usability are subjected to higher support costs, customer dissatisfaction and costs associated with rework [16]. In this subchapter success stories from applications of usability engineering are presented, relevant work published in the context of use studies and different approaches when the intended users are not easily accessible.

- Information Map Studio (IMS) at SAS institute

IMS is an application that enables a technical user (considered an expert user) to create a business view of data, that is relevant to analytical needs of business users which use reporting tools. Project had to overcome many challenges from no established customer base to lack of domain knowledge of BI reporting tools. Initial testing of the original design showed a significant amount of usability issues. After the identified issues were fixed, a second iteration with a follow up usability test was conducted. Results of the test showed that the IMS application had made usability gains on every aspect of the user interface. First iteration of testing included SAS employees who had experience with IMS, but in the second iteration they were excluded in order to minimize the chance of performance improvements due to experience with the product. With the use of a variety of approaches, more value was added to the product teams, simply by combining existing tools and techniques in the ways that best meet the needs of the organization [16] (see pages 112 – 134).
At Nokia, usability tests are used to evaluate the flow of tasks that have been found critical for mobile devices. They faced a challenge with a new product (it was first of its kind on the market) and no real users to conduct the testing. Due to competitive and innovative market, the device was not to be shown to the people outside of the development team. During the development, different methods were used to discover potential problems (usage scenarios, focus groups) and user testing was conducted after the product was released to the public. Users were given the device for some weeks and were asked to report positive and negative features. The results confirmed the developer’s concerns about the effects of consistency with other similar applications that run on desktop machines [37] (see pages 464 – 474).

In both cases, the unavailability of real users was an obstacle. While the approaches in usability testing were different, usability on the products was eventually improved. In this thesis project I faced a similar issue, with a small base of customers, which is not easily accessible for usability activities; therefore I conducted the usability test with proxy users and was able to also perform the test with the actual users of the Customer Care GUI. Results of both tests were similar and during both iterations the same usability issues were discovered. While using proxies has its benefits, there are limitations to consider as well:

- They may know the product or have been involved in the design
- They may experience problems that will not affect the real users (false positives)
- Real users may come across difficulties that do not bother the proxies (false negatives).32

Relation between usability testing and the context of use has been studied within the research communities and there exist a great number of published materials. Research includes a vast variety of applications, ranging from financial sector to health sector.

As mentioned in the introduction, important findings and inspiration for this thesis work is in the work done in the health sector, a study conducted in Norway. Research approached the context of use for specific users of a system and application of the industrial standard was illustrated with three usability studies in their usability laboratory [6]. Particularly of interest are cases A and C presented in the study as they address the context of use in terms of user group and work scenario.

Another case study conducted in the health sector also inspected context of use and usability of a system. Oncology institute was evaluating EHR systems, of whom the main purpose is to handle medical information about patients. Approach in examining the usability was in combination of heuristic evaluation and a survey of experienced users. Results emphasized the importance of the context of use, since there was a significant amount of usability issues found by the novice users. Expert users expressed dissatisfaction with efficiency, flexibility and accessibility of an evaluated system. Authors came to a surprising finding – overall satisfaction was higher than expected, given the dissatisfaction with some aspects of the system [21].

Due to the nature of this thesis work and the area subjected to it, usability and the context of use in publications, this research continues within the electronic and mobile commerce.

Exploratory study conducted to examine usability of general use targeted e-commerce systems (eBay, United Airlines, Yahoo, Amazon) and the conversion of their content onto

32 [http://www.userfocus.co.uk/articles/surrogates.html](http://www.userfocus.co.uk/articles/surrogates.html), (Visited July 2014)
mobile interfaces also emphasized the context of use. Methodology used was a combination of a heuristic evaluation and cognitive walkthrough. As the study suggests, developers of the interfaces should have more understanding when it comes to users and their behavior during the requirement gathering phase. Task subjected to the study were designed with similar steps to their counterparts for PC based access. Novice users had difficulties with completing tasks on a mobile device; therefore their recommendation is to perform usability testing in mobile contexts [22].

A more recent study in usability for mobile devices was addressed in the designing phases of mobile software development and provided guidelines on usable mobile systems. Mobile application ME2.0 (Mobile Electronic Personality version 2) was the subject of examination. Usability evaluation consisted of a pilot user study and interviews, internet based study and prototype. Measurements were acquired through scenarios to assess the user’s comprehension and attitudes regarding the application. Deeper insights on attitudes and expectations of the users were acquired through several iterations. Overall users found the application to be useful and even more if the user interface would simplify its use [23].

Interesting study with a different type of application was also subjected to context of use and usability. Researchers conducted an evaluation of the infotainment system and mobile device in a car. In-car infotainment - IVI (a combination of information and entertainment systems) are becoming common in the newest car models and include a broad range of functions. The context of use in a car was examined with the contextual design method, which is based on contextual inquiry. Observations were made during real journeys that the participants had planned to take regardless of the study. Results of the study contributed to the design and research of future IVI systems based on observational study conducted in a real car context [24].

When dealing with specific products and solutions that are not targeted for mass market context of use is an important factor to consider. In the case of Ericsson’s EWP solution different types of users and the knowledge required to use the Customer Care GUI was measured with the ISO standard [1] that puts the user in focus.

The validity of the proposed method can be maintained in the form of this type of usability testing to be applied to different SA-teams, more test participants and data collected over a longer period of time. This case study was an attempt in that direction.
7 CONCLUSION AND FUTURE WORK

In this chapter I try to summarize the research within usability methods and the knowledge acquired during this thesis project. I return back to research question and try to provide an answer; furthermore I address the subject of the further work I recommend to be done within the EWP solution at Ericsson in regards to usability testing.

A large amount of usability related material was studied in order to examine the methods and possibilities of expansion or modifications. Guidelines that are still valid today are however intended for mass market solutions and World Wide Web interfaces; they target a vast amount of users with different characteristics. Therefore the methods are in need of evaluation for the expert group of users, which are different types of employees that face organization prescribed systems and solutions with various types of interfaces. They are not subjected to using which interface suits their needs most and are by that fact also not exposed to internet trends that target advertising, analytics being collected in order to provide the customized content and advertisements connected to it.

Before answering the research question, I summarize some of the benefits of usability activities within projects in software engineering:

- Design and development identifies the potential usability issues before they undergo coding, which in the long run reduces costs of re-engineering
- No need for a formal usability lab for testing, alternative approaches include remote testing, rooms without recording equipment as long as there is an evaluator present and taking notes, a system that can be accessed from various rooms within organization\(^{33}\)
- Feedback collected from the target group to focus the development and future enhancements
- It increases the usage and the repeat usage
- Minimization of the risk associated with product failing due to poor usability\(^{34}\).

When businesses meet the needs and expectations of their intended users, they are more likely to develop a successful and pleasurable service for their customers.


\(^{34}\) [http://www.experiencesolutions.co.uk/questions/what_is_usability_testing.php](http://www.experiencesolutions.co.uk/questions/what_is_usability_testing.php) (Visited May 2014)
7.1 Answering the research questions

In providing the answers to the research questions, a literature research was conducted as well as the experiment.

- RQ1: How to develop a new M-Commerce usability method which addresses the problem of distance between development and context of use?

With the knowledge obtained from the research communities and usability experts in the field, the proposed method consists of several steps that include:

- test preparation (see section 4.6.1)
- test participant selection – using a team of users that acts as proxies for the actual users (see section 4.6.2)
- task definition – the test cases selected to include the most commonly used daily operation task (see section 4.6.3)
- test execution – simulating the customer solution and configuration (see section 4.6.4).

Proposed method was verified internally and externally and the proposed method also takes into consideration the following:

- expert users of the system
- geographically dispersed expert users with future deployments not known
- development not being in direct contact with the expert users and their potential usability issues.

With the steps described and several factors taken into consideration, the thesis project answers RQ1.

- RQ2: Can proxy users serve as an alternative in usability testing when real users are not easily accessible?

Yes, usability test was conducted with the proxy users and the actual users. Test results of both iterations (see sections 5 and 6.1) were similar and both groups detected the same usability issues.


7.2 Future work

The aim of this thesis project was in examination and evaluation of a specific set of usability guidelines and a specific context of use standard [1], the research area is however much broader than presented and there are other views of usability that can be studied. Of particular interest to the industrial research would be to evaluate other methods related to the ISO standards, such as the SUS questionnaire, which was due to the scope of this project excluded.

Also, the study conducted was targeting the M-commerce area, but the literature suggests other areas are in need of further research of this type.

7.3 Future work in Ericsson

As future work in the usability engineering and the work conducted within User Experience department at Ericsson, this method can be further developed and applied in different parts of the EWP solution, since this thesis work targeted only one interface and one specific group of users. The system overview and the M-commerce ecosystem give an indication of possible areas of usability testing with the proposed method such as:

- Usability testing of other interfaces accessing the system – the expert group of agents conducting their work in the field with their mobile devices. A different group of people would need to participate in that test, proposed are experienced mobile device users in order to measure the context of use for devices other than a PC.

- Usability testing of the other web based interfaces – Admin GUI and Partner GUI, which again employ a different group of expert users in order to detect possible usability problems, here the SA-team´s insight can be used as well.

- Usability testing of installation and deployment procedures. Installation of the system requires a skilled Linux technician in order to deploy software components and connect them to external systems. With the deployment and installation activities being outsourced to countries with low resource cost, these procedures can be verified with the use of integration engineers located in Karlskrona in order to prevent possible problems related to the processes and documentation of the installation procedures.

Listed above are some examples of potential re-use of the proposed method and the possibilities to be expanded beyond the EWP solution can be considered as well.
REFERENCES


46


[23] Oyomno, W., Jäppinen, P., Kerttula, E., Heikkinen, K., Usability study of ME2.0: User interface design for mobile context enhanced personalization software, Personal and Ubiquitous Computing, 02/2013, Volume 17, Issue 2


[38] Cooper A., The inmates are running the asylum: Why high-tech products drive us crazy and how to restore the sanity, ISBN: 9780672316494, 1999

APPENDIX A – SURVEY QUESTIONNAIRE

SURVEY QUESTIONNAIRE

USABILITY/EFFECTIVENESS
1. Where you able to complete this task on existing GUI in the first attempt?
   YES       NO
2. If not-what was the problem? Try to describe what happened? Free text.
3. How often (in your opinion) completing a task takes too long time compared to what you expected?
   Always       Often       Sometimes       Rarely       Never

USABILITY/EFFICIENCY
4. How long did it take you to complete this test?
   0-5 min       6-10 min       11-15 min       16-20 min       above 20 min
5. Did you need assistance during this task?
   YES       NO
6. If yes, how much time did you need to get help?
   0-5 min       6-10 min       11-15 min       16-20 min       above 20 min

USABILITY/ SATISFACTION
7. Do you find the system to be consistent in functionality?
   YES       NO
8. How understandable you find the system messages to be?
   Not at all       very little understandable       understandable
                    quite understandable       very understandable
9. Do you find the interface appealing to you?
   YES       NO
10. If not, what would you like to change to make it more to your liking? Free text

11. To what extent you find the user interface self-explanatory?
   Not at all       very little understandable       understandable
                    quite understandable       very understandable
APPENDIX B – TEST CASES

TEST CASES

- UTC01 - Transfer of funds. An amount of 10 EUR transferred between account holder A and account holder B.

  Objective:
  Test transfer of funds between account holders

  Preparation
  Customer Care user can access Customer Care GUI
  Account Holder A has sufficient funds for transfer
  Customer Care profile allows transfer of funds between Account Holders

  Use Case Execution:
  Step 1: User selects transfer in the “Customer Care GUI” main menu
  Step 2: User selects money transfer
  Step 3: Customer care user enters amount to transfer from Account Holder A to Account Holder B
  Step 4: Customer care user receives notification message
  Step 5: Account Holder A to Account Holder B receives notification messages

  Pass Criteria
  Account Holder A is debited with amount equivalent to the one entered by Customer care. Account Holder B is credited with amount equivalent to the one entered by Customer care

- UTC02 - Edit Account Holder Information - User is to search for an account holder and edit the account holder information – update of the street number in the address field.

  Objective
  Test changing of Account Holder information

  Preparation
  User can access Customer Care GUI
  Customer Care profile allows changing of Account Holder information
  Customer Care can view Account Holder information

  Use Case Execution

  Step 1: Customer care navigates to view account holder information
  Step 2: Customer care select edit account holder Information
  Step 3: Customer care makes changes and confirms changes made on account

  Pass Criteria
  Viewing account holder information from Customer Care GUI shows updated information.

- UTC03 - View transaction history
Objective
Test display of transaction history of account holder

Preparation
User can access Customer Care GUI
Customer care user profile allows view of transaction history
Customer care user can view Account Holder information

Use Case Execution
Step 1: Customer care user navigates to view account holder information
Step 2: Customer care user select view transaction data

Test Case TC1-07-10-01: View historical data
Under view historical data condition Customer Care user has permission to view transaction history of account holder. Initial page shown to the User is search results showing Account Holder details. The following test execution is expected:
0 Customer care user can view account holder information
0 Customer care user select view transaction history

Pass Criteria
Customer Care GUI shows transaction history of selected account holder.

Test Case TC1-07-10-01: View transaction history
Under view vouchers condition Customer Care user has permission to view transaction history of account holder. The following test execution is expected:
0 Customer care user can view account holder information
0 Customer care user select view vouchers

Pass Criteria
Customer Care GUI shows vouchers created by of selected account holder including their status information.

- UTC04 - Eye Principle
Objective
To test four eye principle. The four eye principle is a process that prevents users from performing certain operations in a single step. Instead, an approval is required by another user, for example a supervisor.

Preparation
Customer care profile of user A does not allow approval of changes made on account holder information and transfer of funds between accounts.
Customer care profile of user B allows approval of approval of changes made on account holder information and transfer of funds between accounts.

Use Case Execution – Account Holder Information
Step 1: Customer care user A navigates to view account holder information
Step 2: Customer care user A makes changes on account holder information
Step 3: Customer care user A select request approval of changes made
Step 4: Customer care user B navigates to view approval requests
Step 5: Customer care user B approves Customer care user A changes
Test Case TC1-07-10-01: Account Holder Information

Under account holder information the following test execution is expected:
0 Customer care user A can view account holder information
0 Customer care user A makes changes
0 Customer care user A select request approval
0 Customer care user B can view approval request of Customer care user A
0 Customer care user B approves Customer care user A changes

Pass Criteria
Viewing account holder information from Customer Care GUI shows updated information.
APPENDIX C – ABBREVIATIONS

BI - Business Intelligence
BTH - Blekinge Tekniska Högskola
ECW - Ericsson Converged Wallet
EHR - Electronic Health Record
EWP - Ericsson Wallet Platform
GSMA - Groupe Spécial Mobile Association
GUI - Graphical User Interface
IMS - Information Map Studio
ISO - International Organization for Standardization
IT - Information Technology
IVI - In-Car Infotainment
M - Commerce - Mobile Commerce
MNO - Mobile Network Operator
NDA - Non Disclosure Agreement
PC - Personal Computer
SA - Solution Architect
SUS - System Usability Scale
TAP - Think Aloud Protocol
UTC - Usability Test Case
VDT - Visual Display Terminal