UX METRICS – MEASURING THE SUCCESS OF USABILITY AND USER EXPERIENCE IN LABORATORY AUTOMATION PRODUCT DEVELOPMENT

LAB UNIVERSITY OF APPLIED SCIENCES LTD
Master of Business Administration
Degree programme in International Business Development
Autumn 2020
Petteri Häkkinen

ABSTRACT

Author(s)	Type of publication	Published
Häkkinen, Petteri	Master's thesis	Autumn 2020
	Number of pages 10 + 79 + 11	

Title of publication

UX metrics – measuring the success of usability and user experience in laboratory automation product development

Name of Degree

Master of Business Administration

Abstract

This thesis studied measuring of the outcome and success of usability engineering and user experience (UX) design in laboratory automation product development. The objective was to create a suitable set of UX metrics that could help in the measuring, which could then be used to justify the importance of usability engineering and UX design, and to justify where to allocate the limited design resources.

The theoretical part of the thesis included relevant information related to the theoretical background of the work and the research methodologies that were used. The discussion of the theoretical background introduced the basic concepts such as user centered design, usability, and user experience and included also the principles of UX metrics and the return on investment of UX activities. Lean UX canvas and DMAIC (define, measure, analyze, improve, control) process were selected as the main tools and techniques for the research. The lean UX canvas was selected because it could be used to address the business problem, to find solutions to the problem, and to see towards which direction to move next. DMAIC process on the other hand offered a framework and tools to examine the business problem closer, to analyze the root causes of the business problem, and to find solutions for it.

As the main result of the thesis a recommended list of UX metrics was created. The UX metrics were formed through both the theory introduced in the thesis and the experiments that were conducted to collect data and to ensure that the UX metrics work in their intended context. The thesis succeeded in creating a set of UX metrics that can be recommended to be used when measuring the outcome and success of usability engineering and user experience design in laboratory automation product development.

Keywords

DMAIC, Laboratory automation, Return on investment, Usability, User experience, UX metrics

PREFACE

A couple of years ago, I was in a situation where I wanted to acquire something new professionally. Back then I decided to educate myself in the area of International Business Development and applied for a Master of Business Administration degree. Now two and half years later this journey has now come to an end.

This thesis is made during 2020 to examine the topic of UX metrics and to support the usability engineering and user experience design at Thermo Fisher Scientific. I want to thank Thermo Fisher Scientific for offering me the possibility to combine the topic of my thesis with my work tasks and for offering me the access to the data used in this thesis.

In this thesis I decided to use DMAIC process, one of the techniques in Lean Six Sigma. I have Lean Six Sigma Yellow Belt myself and although leading a DMAIC process is normally only done by Green and Black Belts, I decided to give it a try. After all, there was no project team involved in this process. How successful it was, I leave for the readers to decide. For me it acted as a good learning exercise.

For guidance on the journey of this thesis I want to thank Brett Fifield, Principal Lecturer in International Business. Our monthly conversations gave me always new thoughts regarding the topic and kept me on track with the thesis.

I would like to show my appreciation to my parents and family for their continued support and encouragement throughout the years. Without you and your encouragement to make my own choices, I would not have ended up exactly where I am at the moment. I have learned a lot from you.

I would also like to express my special thanks to my wife, Marika, who I know will always be my support no matter what happens. You encourage me to be the person I am. A big thank you also to our dog Lulu, who kept me company while I was working long hours with the thesis.

Finally, I would like to thank everyone who has helped and contributed to the creation of this thesis, as well as supported me at various stages of the studies.

Helsinki, 1 December 2020

Petteri Häkkinen

CONTENTS

LIST OF FIGURES

LIST OF TABLES

TERMS AND ABBREVIATIONS

CSAT

Customer satisfaction.

CTQ

Critical-to-quality.

DOE

Design of Experiments.

DMAIC

Define, Measure, Analyze, Improve, Control.

Formative usability evaluation

The main target of formative usability evaluation is to explore the strengths, weaknesses, and use errors of the user interface design. The formative usability evaluation activities are generally conducted iteratively throughout the whole product development process. (IEC 62366-1 2015, 9.)

Human-centred design

Another commonly used term for user centered design. See user centered design for description.

ISO

International Organization for Standardization.

IVD

In Vitro Diagnostics.

KPI

Key Performance Indicator.

NPS

Net Promoter Score.

ROI

Return on Investment (ROI) is about evaluating how efficient an investment is (Chen 2020). It is a performance measure of what is got back compared to what is put in (Duermyer 2020). When calculating ROI, the return of the investment is divided by the cost of the investment. The result of this calculation is presented as a ratio or a percentage. (Chen 2020.)

SIPOC

Suppliers, Inputs, Process, Outputs, Customers.

Summative usability evaluation

Summative usability evaluation is performed at the end of the product development. It is the last evaluation performed for a product and its goal is to obtain objective evidence that the use of the device can be performed safely without use errors. (IEC 62366-1 2015, 10.) Summative usability evaluation can also be called as usability validation.

SUS

System Usability Scale.

SWOT analysis

SWOT analysis is a framework that lists the strengths, weaknesses, opportunities, and threats of the object that is under analysis (Hill & Westbrook 1997, 46).

Usability

Usability defines how easy a user interface is to use. It consists of five quality components that are learnability, efficiency, memorability, errors, and satisfaction. (Nielsen 2012.)

With medical devices IEC 62366-1 (2015, 21) defines usability to consist only of safety and efficiency. It means that the users of medical devices must be able to perform the needed tasks safely and efficiently.

User

End user who uses the laboratory automation system. Typically a laboratory analyst or laboratory technician.

User centered design

Approach to design and development that aims to make products more usable by focusing on the use of the product and applying usability knowledge and methods (ISO 9241-210 2019, 2).

UUA

Utility, usability, and aesthetics.

UX

User experience. According to ISO 9241-210 (2019, 4) standard user experience is defined as:

> *User's perceptions and responses that result from the use and/or anticipated use of a system, product or service.*

One could also say that user experience is the feelings you get when interacting with a device in front of you in the moment you are using it (Derome 2015).

UX KPI

User experience Key Performance Indicator. In some occasions UX KPIs are used as a synonym to UX metrics.

UX maturity

UX maturity defines the level how well user experience is taken into consideration in a company (Plewes & Fraser 2015).

UX metrics

UX metrics are signals or attributes that show whether the chosen user experience strategy is working or not. With the help of UX metrics changes can be tracked over time. (Pavliscak 2014.)

VSM

Value Stream Mapping.

VOC

Voice of the Customer.

1 INTRODUCTION

Too often people feel themselves confused or even frustrated when using technical devices. They make mistakes and the devices are not functioning in a way they had hoped for. When thinking about the reasons behind these use mistakes, too often the argument that it was just a user error is heard. The mistakes made are not anyway fault of the users. According to Donald A. Norman (2013, 7) when users make mistakes it is because the products are designed in a way that causes them to be used wrongly.

When the product does not work logically it is easy to blame the user from the errors. With the right type of design, it is anyway possible to remove these errors or at least decrease the amount of them. When concentrating on the users already in the design phase, the product features can be more usable and the use of the product more satisfying. In order to fulfill the real needs and requirements of the users, it is important to put effort to the design of user interfaces.

The first chapter of this thesis explains the background for the thesis and gives a reasoning why the topic has been chosen. It also lists the objectives of the research and the research questions that guide the research process. Lastly a quick look of the structure of the thesis is given.

1.1 Background

Many organizations use a significant amount of time in improving their products. They might ask themselves the following questions: How to improve the attractiveness of the product to customers? How to modify the current versions of the products? How to expand the customer base? (Juti 2016, 139.)

Although competition may be tough the answer to these questions does not lie in copying from the competitors. In copying others, one does not understand why certain decisions were made. Development does not happen without understanding. The organizations who copy are passive and always one step behind others. Their products are mainly weaker versions of the originals. (Fried & Heinemeier Hansson 2011, 135-136.)

Instead of copying organizations should aim to create products that are more than just basic goods. They should be something that competitors are not able to offer. (Fried & Heinemeier Hansson 2011, 138.) Investing in usability engineering and user experience (UX) design can offer these possibilities to differ from competitors.

Investing in usability engineering and user experience design brings benefits but it can also be time consuming and expensive. Organizations do not want to spend money on wrong things. Spending money on wrong things drives down margins and does not buy the time to achieve sustained sales (Rosser 2009, 31). For this reason it is important to know where to concentrate in the design. At some point the design is satisfying the user needs and extra investments after that point should be aimed to user experience. This has been proven already decades ago as shown in Figure 1.

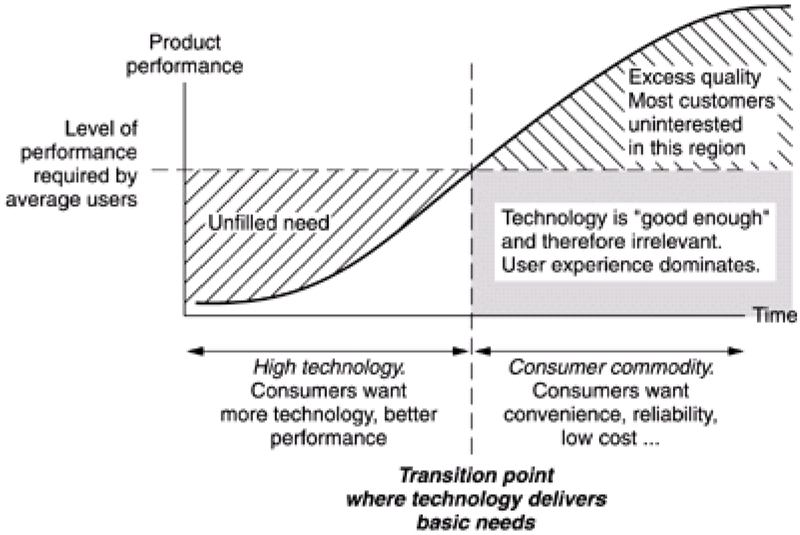


Figure 1: Needs-satisfaction curve of a technology (Norman 1998).

This thesis is done to study the usability engineering and user experience design activities in laboratory automation device design and development at Thermo Fisher Scientific. The final results of the thesis are intended to be used in the daily work of the UX team and the intended readers of the thesis are primarily the members of the UX team, research and development personnel, and product management personnel. Laboratory automation devices that are in the target of the thesis are medical devices or more specifically IVD (in vitro diagnostics) devices that perform different kind of pre and post analytical tasks when handling sample tubes (Thermo Fisher Scientific 2020). The users of the devices are typically laboratory analysts and laboratory technicians (Thermo Fisher Scientific 2019).

The laboratory automation devices form a system that automates many tasks (e.g. removing caps from sample tubes, inserting caps to sample tubes, and dividing sample from primary tube into several secondary tubes) that have previously been performed manually in the laboratories. This changes the nature of the work tasks for the personnel in the laboratories. What used to be a lot of manual work has changed to be taking care of that the laboratory automation system works as expected. The work tasks with the laboratory automation system include e.g. adding consumables to the system, performing user maintenance tasks, and solving possible problem situations in order to keep the system running. Here the usability and user experience of the devices stand out. (Thermo Fisher Scientific 2019.)

According to studies in Finland as much as 8 % of total work hours can go to problem solving of technical devices (Wiio 2004, 18). Reports from IBM show similar results but they also claimed 100 to 1 return on investment in usability testing because of reduced training, maintenance, production, and revision costs (Shneiderman 2003, 26). In general the results have been noticed also in the product development of laboratory automation devices.

The status and appreciation of usability and user experience seem to have grown in the past years. It feels that more and more value is given to those activities in the laboratory automation product development. In some sense the users are taking usability already for granted or at least demanding usable devices. This applies both to the software and hardware of the devices.

For its part this thesis is trying to respond to the same theme. Focusing the limited resources in use to right things can be challenging and not always that simple but with right actions it can be done.

1.2 Research questions

In short the objective of the thesis is to study measuring usability and user experience, and to create a list of UX metrics that can be used to guide the usability engineering and user experience design activities when developing laboratory automation devices. Utilization of the UX metrics in the measurements can help bringing extra value for the business. Utilizing them in the measurements and then using the results of the measurements can e.g. help focusing the limited design resources on right things, show the benefits of usability engineering and UX design, increase customer satisfaction, improve brand image, bring extra value to the end users, and achieve savings.

The theory part of the thesis collects information related to UX metrics and ROI of UX activities. The empirical study part of the thesis explores and evaluates possibilities of UX metrics. The aim is to use both theory and the results of the empirical study to find UX metrics that are suitable and work well in the development of laboratory automation devices.

This thesis has four research questions in total (see Figure 2). The main question is related to UX metrics. In addition to that there are three sub-questions that are tightly tied to the main question. The sub-questions are related to return on investment of UX activities, internal and external stakeholder information, and future utilization of the metrics. The thesis aims to find answers to all four of the research questions.

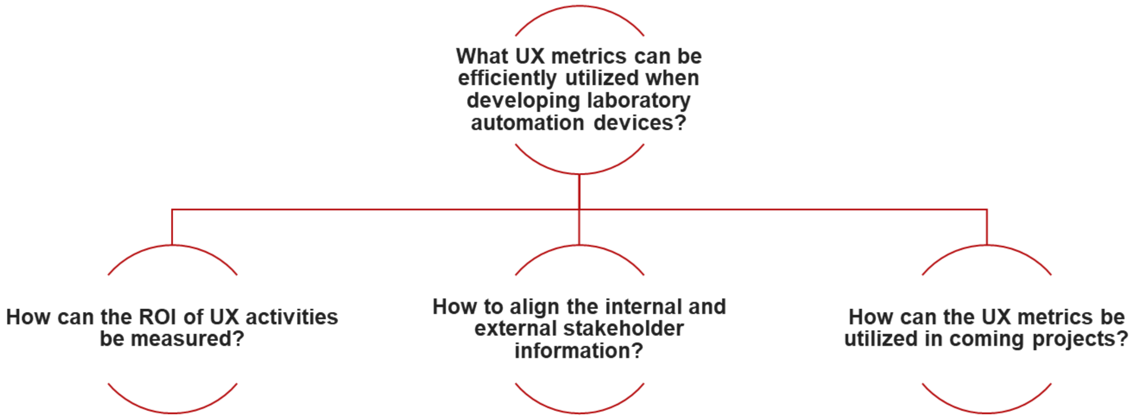


Figure 2: Research questions.

UX metrics: What UX metrics can be efficiently utilized when developing laboratory automation devices?

UX metrics are commonly used to measure the user experience and to communicate the business impact of the UX design. There is not anyway a set of metrics available that would be proven to work in laboratory automation device product development. Finding out suitable UX metrics to be used when developing laboratory automation devices will help figuring out where to concentrate in the UX design. The UX metrics will help to put the focus on things that bring more value.

Return on investment: How can the ROI of UX activities be measured?

Putting a price tag to UX activities is not simple and even more challenging is to find out the ROI of UX activities. There are anyway many possibilities that can offer return on investment. Finding out the ROI of UX activities will help concentrating the resources to activities that bring also monetary value.

Stakeholder information: How to align the internal and external stakeholder information?

Multiple stakeholders are involved in the design and development of the laboratory automation devices. They all have also their own needs and requirements for the end products. Aligning the internal and external stakeholder information becomes important when design resources are limited. Internal stakeholders include for example people working in research and development, product management, and production. External stakeholders on the other hand consist for example of end users, distributors, and laboratory management.

Future: How can the UX metrics be utilized in coming projects?

In order to help bringing extra value, the UX metrics must be taken into account in the future projects. Otherwise creating them will be useless. When creating the UX metrics it must be kept in mind that they need to be suitable to work with laboratory automation devices.

1.3 Structure of the thesis

This thesis is divided into seven chapters. The first chapter acts as an introduction by explaining the background and the research questions. Chapter two presents the starting point of the thesis by defining the problem space and the current situation of the usability engineering and user experience design. Together the first two chapters define the purpose of the research and form a hypothesis for the research.

Chapters three and four form the theoretical basis for the thesis. They introduce the supporting theories, relevant literature, and methodologies. Chapter three discusses about relevant theoretical background and terms concentrating on user centered design, usability, and user experience. Chapter four on the other hand explains the theory behind research methodologies that are used in the thesis.

Chapters five and six form the empirical part of the thesis. In chapter five the performed research, data collection, design of experiments, and analysis activities are explained and in chapter six the research findings and recommendations are presented. Chapter six gives also answers to the research questions.

Lastly chapter seven gives a conclusion for the thesis. It also contains an analysis how the thesis managed to fulfill its objectives and discusses about future recommendations.

2 STARTING POINT OF THE RESEARCH

The second chapter introduces the starting point for the research performed in this thesis. The chapter includes a description of the problem space and of the current situation of usability engineering and user experience design activities.

2.1 Problem space

The research in this thesis is done by studying the usability engineering and user experience design in laboratory automation device product development at Thermo Fisher Scientific. The thesis is done to support the usability engineering and user experience design activities, to help realizing where to concentrate the limited resources in use, and to figure out if monetary value can be given to the outcomes of the activities directly.

The research is limited to the laboratory automation devices of Thermo Fisher Scientific only. The devices can be individual modules or whole laboratory automation systems consisting of modules. The systems can be large and consist of several modules or rather small and consist only of few modules. (Thermo Fisher Scientific 2019.) In any case the laboratory automation devices of other manufacturers are not taken into consideration.

Currently the problem is that the success of usability engineering and user experience design is not measured as effectively as it could be done. There are no concrete tools that could be used to perform the measurements. The resources in use are limited and every area cannot be invested equally. Neither does it make sense to invest UX design resources into every area to the same extent. It also feels even today that the usefulness and benefits of usability engineering and user experience design have to be justified to some people in the organization. (UX team 2020.)

Some areas, that are critical to quality from customer's point of view, are a functioning device, safety of use, effectiveness of use, and usability (Thermo Fisher Scientific 2019). Keeping these in mind a suitable set of user experience metrics to measure the success of the usability engineering and user experience design activities should be created. Most likely the metrics contain both monetary and soft skill metrics, and information from internal and external stakeholders, trying, and adjusting are needed to create them.

Where to focus the resources in the usability engineering and user experience design processes needs clarification and the key here are functioning user experience metrics. They will help deciding in which areas to focus in the future and where to allocate the resources.

2.2 Current situation

The current usability engineering and user experience design activities are already in good shape at Thermo Fisher Scientific. Already for a few years there has been an own in-house user experience team in the research and development department. The team consists of people with various backgrounds, contains the necessary know-how, and is able to bring in user experience expertise to projects as quickly as needed. (UX team 2020.)

There is a usability engineering process in place and it is integrated with other product development processes. This ensures user experience team's involvement and appropriate techniques to be used in all phases of the projects when needed. The way of working is project based and there are several projects going on at the same time. Several ongoing projects keep the people busy so there is little time for further development of the processes. (UX team 2020.)

Currently the success of the usability engineering and user experience design is measured mainly in the summative usability evaluation performed for the devices (UX team 2020.). There are only two metrics that are followed: safety and effectiveness. Safety means that during the evaluation there are no hazardous situations that could cause harm to the user, device, or environment. Effectiveness on the other hand means that users are able to perform all required tasks with the device. (IEC 62366-1 2015, 21.) Measuring both safety and effectiveness comes from the IEC 62366-1 (2015) standard that is followed in the usability engineering activities. To comply with the standard measuring safety and effectiveness is mandatory.

The UX team's resources available for new product development projects are limited and every area cannot be invested equally. All project teams have members from the UX team but it is necessary to know in which areas especially to concentrate in the laboratory automation device's usability and user experience. When the success of certain activities can be measured with the help of UX metrics, it will become easier to allocate the UX design resources to certain areas based on the results of the UX metrics.

Although the in-house user experience team has been existing for a few years, the usability engineering process is integrated with other product development processes, and the usability engineering and user experience design activities are part of every product development project, there are still people in the organization who think that user experience is not needed and that the UX team is only bringing more work to other teams in the research and development department. Sometimes it feels that still after years the usefulness of usability and user experience needs to be justified to certain people. (UX team 2020.) When the success of usability and user experience can be properly measured with the help of UX metrics, also the opinions of these people could be influenced.

3 LITERATURE REVIEW

According to Donald A. Norman (2016, 5) technology is:

New stuff that doesn't work very well or that works in mysterious, unknown ways.

Norman is also amazed how people are willing to spend weeks, months, or even years to learn how to drive a car or how to play an instrument, but at the same time complain if they have to spend 15 minutes to learn how to use new technologies. Luckily, it is possible to have an effect on how to use technologies by investing time and effort in the user interface design. When investing in the user interface design it is possible to improve the usability and user experience of the technologies and to decrease the time needed to learn how to use the technologies.

This theoretical literature review consists of three main parts that explain the theoretical background and terms that are relevant in the usability engineering and user experience design of laboratory automation modules. In the chapter basic principles of user centered design, usability, and user experience are introduced.

The first part of the literature review concentrates on user centered design. User centered design activities are the basis for usability engineering and user experience design in laboratory automation product development. Understanding the basics of user centered design help getting a better picture of the process. The second part is about usability. The basic concept of usability is guiding the design work. In today's world the users of the laboratory automation devices are already expecting the devices to be easy to use and work as intended. The third part focuses on user experience. Level of UX maturity in a company, the use of UX metrics, and the ROI of UX activities are explained to provide understanding of the key topics of the thesis.

The objective of the theoretical literature review is to give an understanding of the basics of user centered design, and explain the concepts of usability and user experience. All these three elements are essential for the research performed and the recommendations provided later in this thesis.

3.1 User centered design

The basis of user centered design is a systematic collection of user information (Hyysalo 2009, 77). In addition to performing the user research, also knowledge about usability in general and various user centered design methods are important. The International Organization for Standardization (ISO) defines user centered design in ISO 9241-210 (2019, 2) standard in the following way:

Approach to systems design and development that aims to make interactive systems more usable by focusing on the use of the system and applying human factors/ergonomics and usability knowledge and techniques.

In practice this means according to ISO 9241-210 (2019, 6) that user centered design should always obey certain principles. The design is based upon an explicit understanding of users, tasks, and environments. The design is also driven and refined by user centered evaluation and it addresses the whole user experience. Users should be involved in the process throughout the design and development and it is important that the process is iterative. The design team should also consist of people with various backgrounds in order to include multidisciplinary skills and perspectives. (ISO 9241-210 2019, 6.)

3.1.1 User centered design activities

After identifying the need for developing a product and using user centered design process for this, understanding and specifying the context of use, specifying the user requirements, producing design solutions, and evaluating the design shall take place in the design and development of the product (ISO 9241-210 2019, 10). The interdependence of these four activities is shown in Figure 3.

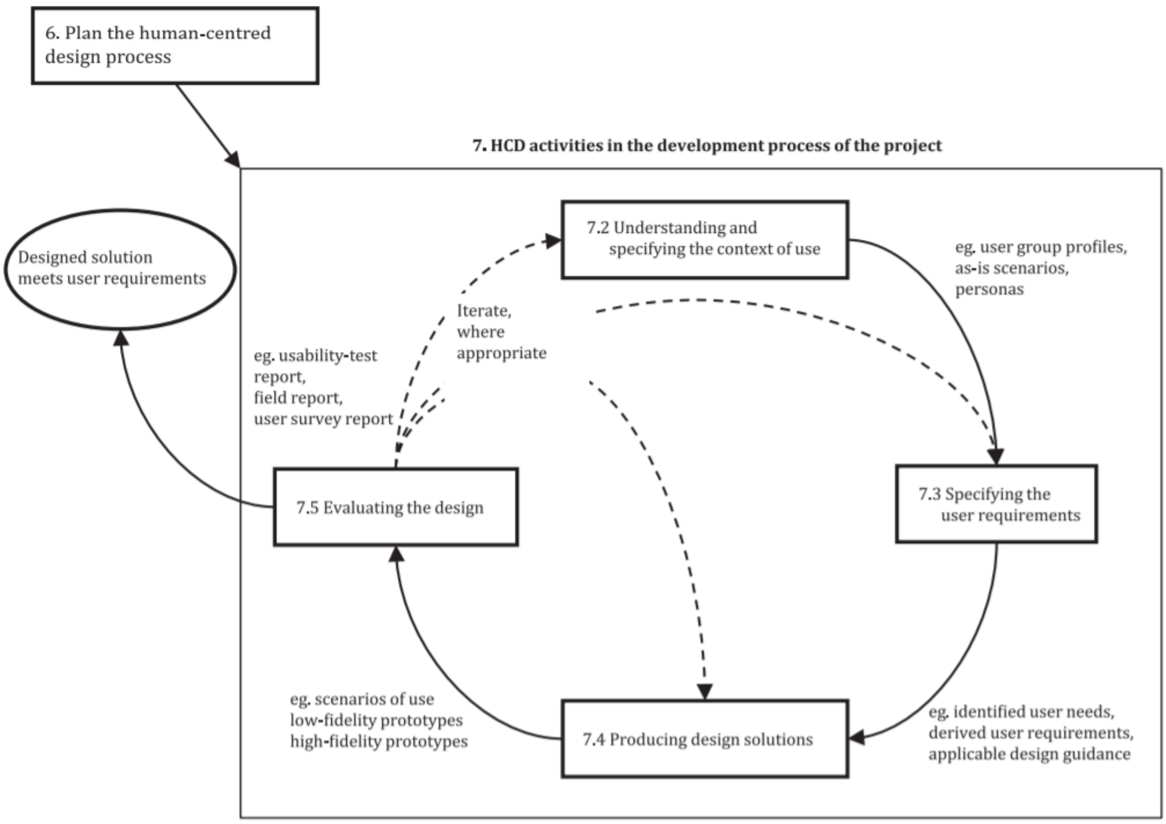


Figure 3: Interdependence of user centered design activities (ISO 9241-210 2019, 12).

The first activity is to understand and specify the context of use. First of all the context of use description should describe who the users of the product are. A product can have multiple user groups who have different needs for the product. From the users it is good to understand who the users are, what their needs are, how they are currently trying to perform their tasks, and how they would like to perform their tasks (Marsh 2018, 3). In addition to user characteristics (knowledge, education, skill, experience etc.) the goals and tasks of the user play a significant role in the context of use description. The context of use includes also description of the use environment. The use environment includes the technical, physical, social, and cultural environments. (ISO 9241-210 2019, 13.) The context of use must be described detailed enough so that it will support the following activities.

Specifying the user requirements creates a basis for fulfilling the real needs of the users. The specification of user requirements includes the intended context of use, requirements derived from user needs, requirements arising from relevant ergonomics and user interface standards and guidelines, usability requirements and objectives, and requirements derived from organizational requirements that directly have an effect on the user. It is important that the user requirements are defined in a way that allows testing them in later phases. (ISO 9241-210 2019, 13-14.)

The producing of design solutions is made based on the context of use and user requirements. The design decisions will have a major impact on the user experience of the product. The design solutions should be made concrete by using scenarios, simulations, prototypes, and mock-ups, and when needed they should be modified in response to user centered design evaluation and feedback. (ISO 9241-210 2019, 15.)

Evaluating the design is performed as user centered evaluation. It means that the evaluation is based on user's perspective. The evaluation should happen in all phases of the product development. Two widely used approaches to user centered evaluation are user based testing and inspection based evaluation that uses usability guidelines. In user based testing the users carry out tasks using the actual product, prototype of it, or by going through models and scenarios. Inspection based evaluation on the other hand is performed by usability experts. It is simpler and quicker to carry out than user testing but does not always find the same problems that would be found in user based testing. (ISO 9241-210 2019, 17-19.)

3.1.2 Arguments for and against user centered design

During its existence there have been various arguments for and against user centered design. For example ISO 9241-210 (2019, 5) describes substantial economic and social benefits as a rationale for adopting user centered design methods. According to the standard the overall quality of products designed using user centered design methods is improved by increasing for example the usability of products (effectiveness, efficiency, and satisfaction), the accessibility of products (for people from a population with the widest range of user needs, characteristics, and capabilities), the productivity of users by reducing discomfort and stress, and the operational efficiency of organizations. The user experience of the products can be improved and the products will be easier to understand and use, which will thereby reduce the training and support costs. Using user centered design methods can also provide a competitive advantage e.g. by improving brand image and contributing towards sustainability objectives. (ISO 9241-210 2019, 5.)

Despite of the clear advantages of user centered design there are also arguments against it. Usually they include the same arguments that user centered design is expensive and although user centered design methods were used the users do not think that the product is significantly better than competitors' products. Göransson, Gulliksen, and Boivie (2003, 115-118) have discussed about arguments against user centered design.

Often as an argument against user centered design it is said that involving users is expensive and of little value since user centered design requires an up-front investment that is tied to uncertain returns. According to the arguments studying users is not useful since user studies can easily confuse what users want with what they truly need. User centered design techniques are also said not to replace good design. Creating usability requirements on the other hand can be seen useless since they are anyway the first ones to cast aside when time is running out in the development project. (Göransson, Gulliksen & Boivie 2003, 115-118.)

In general nowadays large international companies who produce laboratory automation devices (such as Thermo Fisher Scientific) seem to have understood the value and benefits of user centered design. The companies are willing to make the up-front investment and trust that it pays back e.g. in the form of reduced training and support costs or by improved brand image. Unfortunately it is also seen that when the time is running out in the development projects often the usability requirements are the first ones that will be compromised.

3.1.3 Four waves of user centered design

In the ever-changing world also user centered design has had to adapt to the constant change and take new needs into consideration. The basic techniques and methods have remained the same throughout the years but the focus has changed according to the changes in the world.

Gribbons (2013) lists three guiding principles that have remained constant although much has changed over the years. First of all user centered design begins and ends with a deep and comprehensive study of human behavior related to product and experience design. Secondly, unless value is delivered to business, there will never be the opportunity to serve the user. Thirdly, strategies and research methods underlying user centered design offer valuable contributions to much broader applications in any place where a product or service is designed for use by people. (Gribbons 2013.)

In the past few decades four waves of user centered design can be recognized (see Figure 4). In the first wave user centered design focused on external support, i.e. on documentation, training, and customer assistance. It can be clearly seen that the focus was on solving the technical aspects of product design. After years of focusing on external support the focus on user centered design moved more into usability. (Gribbons 2013.) This could clearly be linked to time when computers began to spread to households and user friendliness became a goal for home applications (Kolko 2018, 31). The second wave was guided by a definition of usability that included ease of learning, ease of use, transfer of learning, and segmentation of users. Technology was not seen as an end, but rather as an enabler, supporting the achievement of real user goals and serving real human needs. (Gribbons 2013.)

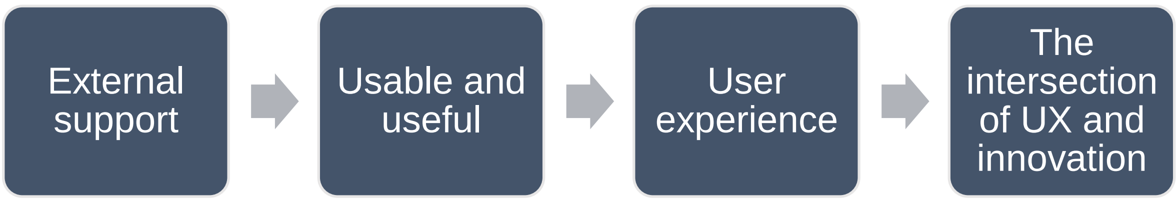


Figure 4: The four waves of user centered design according to Gribbons (2013).

According to Gribbons (2013) the third wave was a push towards user experience caused by new technologies like World Wide Web and mobile technology. Also notable was an integration of user experience in business strategy and development. User experience became part of a product's brand and was carefully orchestrated across every touch point with the customer. (Gribbons 2013.) The change led to companies using the term user experience as a key differentiator in the competitive markets (Pacheco 2019).

The fourth and the most recent wave is the intersection of UX and innovation which is guided by a focus of real and unmet human needs. After the stable waves of usability and user experience suddenly there was an opportunity for continuous product development where each new product or enhancement is quickly followed by another. (Gribbons 2013.)

When thinking about the product development of laboratory automation devices and their relationship to the evolution of user centered design, it can be noticed that medical devices are not following the latest trends. The product development takes a long time, the release rate is not very fast, and the standard guiding the work (IEC 62366-1 (2015)) is mainly concentrating on usability. Laboratory automation devices being usable and useful is already taken for granted but user experience and innovativeness are factors that might offer points of differentiation in the competitive markets in the future.

3.2 Usability

Usability is a quality that many products possess but many more lack (Rubin & Chisnell 2008, 3). Often people tend to pay attention to usability when a product is not functioning as assumed. This causes dissatisfaction. When usability is on a good level it is not even noticed since the products are functioning as they should.

Bad usability can cause problems also to others than just to the people using the products. As an example if employee's time is wasted due to bad usability it can mean also financial losses to the employer or as Nielsen (2000, 14) puts it, bad usability equals no customers.

3.2.1 Definitions

ISO 9241-11 standard (2018, 11) defines usability in the following way:

> *Usability is the extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use.*

The definition defines three components of usability. Effectiveness is the accuracy and completeness with which users achieve their specified goals (ISO 9241-11 2018, 14). Efficiency means the resources that are used in relation to the results achieved (ISO 9241-11 2018, 15). Satisfaction on the other hand means the extent to which the user's physical, cognitive, and emotional responses meet user's needs and expectations when the system, product or service is used (ISO 9241-11 2018, 16). Figure 5 presents the three components of usability and their relationship with outcomes of use and context of use.

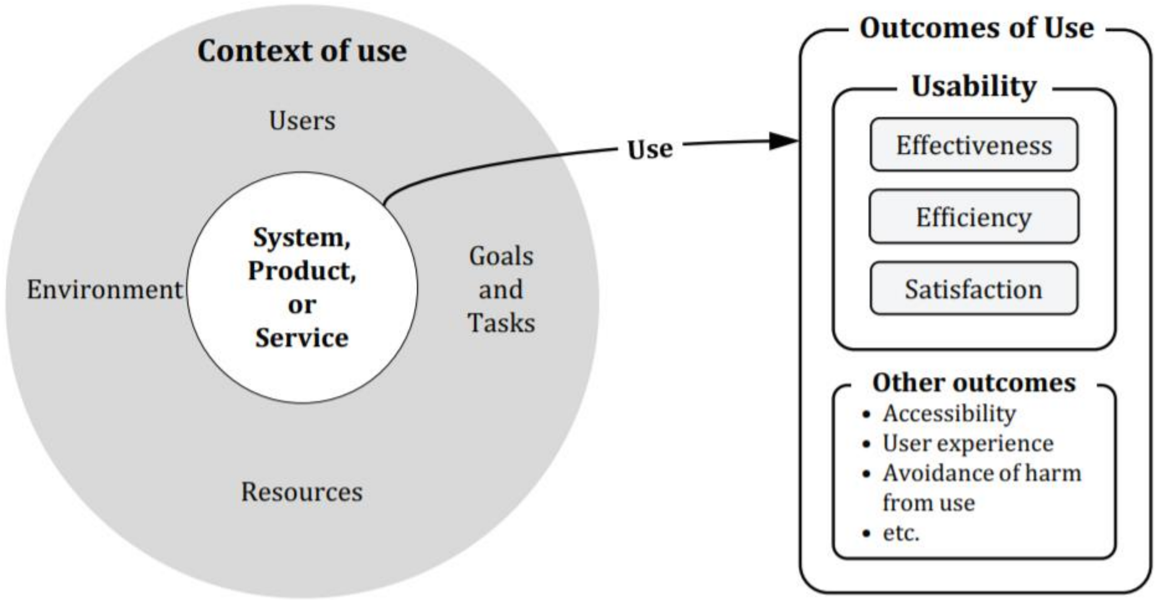


Figure 5: Usability that results from use of a system, product, or service in a context of use (ISO 9241-11 2018, 12).

Probably the most commonly used definition for usability comes from Jakob Nielsen. Although being almost thirty years old, the definition is still valid and cited in publications all over the world. Nielsen divides usability into five components which are learnability, efficiency of use, memorability, errors, and satisfaction (Nielsen 1993, 25).

Learnability is probably the most essential component of usability. Using products must be easy to learn since the first experience most people have with products is to learn how to use them. Highly learnable products allow users to reach a reasonable level of usage proficiency within a short period of time. (Nielsen 1993, 27-28.)

Efficiency of use refers to expert user’s steady state level of performance at the time when their learning curve flattens out (Nielsen 1993, 30). It is about how quickly and productively the users are able to perform the tasks once they have learned the design (Nielsen 2012).

Memorability means that casual users do not have to learn the use of the products again after a period of not using them. Instead of learning the users will remember how the products work based on their previous learning and this way they will re-establish their proficiency. (Nielsen 2012.)

Errors as a component of usability means that users should make as few errors as possible when using the product. There should also be no catastrophic errors at all. If some errors anyway happen, recovering from them should be as easy as possible. (Nielsen 2012.)

Satisfaction is describing how pleasant the use of the product is. A product should be so pleasant to use that users feel themselves satisfied during the use of the product. (Nielsen 2012.)

Antti Wiio on the other hand defines a usable system to be understandable, easy, comprehensive, and aesthetically pleasant. When the system is understandable users are able to figure out how the system works and how to reach the wanted outcome. Easy system lets the user complete the needed tasks in a simple and fast way. Comprehensive system offers all the features and information that are needed to handle the tasks. Aesthetically pleasant system gives a signal of quality and know-how. (Wiio 2004, 14-16.)

The definitions of usability defined in ISO 9241-11:2018 standard and by Nielsen and Wiio are all broader than the one used in general with medical devices. IEC 62366-1 (2015, 21) defines medical device usability to consist only of two aspects: safety and efficiency. It means that users must be able to perform the needed tasks safely and efficiently.

3.2.2 Why are products difficult to use?

A web store where you cannot find the information you are looking for. A remote control that is full of unnecessary buttons. A washing machine that requires reading the user manual every time you want to turn it on. There are numerous examples of bad usability and products that are difficult to use. According to Shneiderman (2003, 141) bad products are produced as long as consumers are willing to do business with companies that produce them. Luckily nowadays it is relatively easy to explore other companies and to see their selection of products too.

Many reasons can lead to bad usability that causes products to be difficult to use. Rubin and Chisnell (2008, 6) have recognized five reasons for the development of products that are difficult to use. These reasons are that development focuses only on the machine or system, target audiences change and adapt, designing usable products is difficult, team specialists do not always work in integrated ways, and that design and implementation do not always match (Rubin & Chisnell 2008, 6).

Since products often require technical solutions to be implemented, the focus is easily put to the machine or system instead of the users (Wiio 2004, 53). As an argument for this it is often said that humans are so adaptable that it is easier to let them adapt themselves to the product requirements than vice versa. Typically developers who create the products have been hired because of their skill to solve technical problems. For them communicating with the users means stepping out of their own comfort zone. (Rubin & Chisnell 2008, 7.) Information regarding users and context of use are anyway essential for product design (Hyysalo 2009, 63). If product designers do not know the users it is difficult to match the product to the user requirements. This emphasizes the importance of user research also in laboratory automation product development.

Back in the days especially the users of computer based products were experts in the area and developed the systems by themselves. Usability problems did not matter because the users were happy to troubleshoot and repair all the problems. Nowadays this is not the case anymore. The user groups have expanded and continue to change dramatically. The users do not possess similar knowledge as the developers and it is expected that the products work without the constant need of troubleshooting and repairing. (Rubin & Chisnell 2008, 8.)

Designing usable products is difficult but still many organizations tend to think that just using common sense is enough. This way of thinking brings many problems that can be dangerous what comes to usability. In the organizations everyone seems to have an opinion about usability and how to achieve it. When it comes to the point that it is time to evaluate the usability of a product they anyway do not have the required skills to do that. (Rubin & Chisnell 2008, 9.) The amount of designers who have never met or talked to users is alarming (Shneiderman 2003, 32). Designing usable products is challenging without communication to the intended users but it is even more challenging if the designers do not recognize the factors that affect usability.

Design and development organizations often consist of groups of experts who have their own ideas and approaches what comes to product development. Unfortunately too often they fail to integrate themselves with each other which can be seen in the end product. As a solution to improve efficiency, many organizations have broken down the product development process into separate components that are developed independently by separate teams or individuals. Here the problems arise because of poor communication among the teams. When looking from outside it looks like the teams are running different projects in their own silos. The users are anyway expecting that the product works as a whole no matter how it is designed. Only if the components work well together, the product will be viewed as usable and it can meet the user requirements. (Rubin & Chisnell 2008, 9-10.)

The design and technical implementation of a product are two separate activities that require very different skills. Due to expanding user groups especially the need for good design skills has grown. (Rubin & Chisnell 2008, 11.) It is desirable that design and development organizations ensure that all the necessary knowledge and skills are found within the organization.

3.2.3 What makes products more usable?

It is not always easy to find an answer to the question of what makes a product more usable. There are anyway many ways to affect the usability and probably the most important is to concentrate on the users. Rubin and Chisnell (2008, 13) emphasize the importance user centric design and especially the three basic principles of it: early focus on users and their tasks, evaluation and measurement of product usage, and iterative design.

In addition to just simply identifying and categorizing the users, a direct contact between the users and the design team throughout the development life cycle is recommended. The product should be evaluated with actual users throughout the whole development process. Based on the evaluation it must be possible to even completely rethink the design, which is the core of true iterative design. (Rubin & Chisnell 2008, 13-14.)

Göransson, Gulliksen, and Boivie (2003, 115-116) have done research regarding how to increase the usability in software development. Based on their research they recommend several conclusions how to increase the level of usability. First of all usability tools, techniques, and methods must be integrated in and relate to the software development process. Usability professionals must learn more about software development and gain a better understanding of the possibilities and limitations of software development tools. At the same time also software developers must learn more about usability and user centered design. User involvement is crucial and the development process must allow for the time it takes. Also enough space and time must be allowed for interaction design activities. Attention must be paid to the use of prototypes and scenarios as well as to the work practices of the multidisciplinary team. (Göransson, Gulliksen & Boivie 2003, 115-116.)

In laboratory automation devices the software development has a big role since the devices contain both hardware and software With the software it is easier to make also modifications to existing devices that are already out in the field. In order to increase the usability in software development a lot of cooperation is done. Attention is paid to the work practices, and usability professionals and software developers try to learn from each other’s work.

3.2.4 Importance of usability

Usability is important for several reasons. It is an important factor for example when marketing products. Nowadays the competition is tough in almost all fields. Users do not buy just anything but they require more and more from the usability of the product. (Kuutti 2003, 15.)

On the web, usability is a necessary condition for survival. If a website is difficult to use, users leave the site. There is no such thing as a user reading a website manual or otherwise spending much time trying to figure it out. There are plenty of other websites available and leaving the site is the first thing users do when they encounter a difficulty. On ecommerce the first law is that if users cannot find the product, they cannot buy it either. For company internal intranets, usability is a matter of employee productivity. (Nielsen, 2012.) If time is wasted because of usability problems it equals paying employees without getting the work done (Kuutti 2003, 16).

According to Nielsen (2012) best practices call for spending about 10% of a design project's budget on usability. This has clear advantages. On average it will more than double a website's desired quality metrics and slightly less than double an intranet's quality metrics. For software and physical products, the improvements are typically a bit smaller but can still be considered substantial.

Usability has also an important effect what comes to product safety. Conducting errors with medical devices or in the cockpit of an airplane can be fatal. When not being fatal usability problems can anyway be frustrating and this can lead to risks of accidents at work. (Kuutti 2003, 16.)

3.3 User experience

Although in medical device design usability has the main role, user experience (UX) is significantly involved in the design activities. As seen in the four waves of user centered design the focus in general has changed from usability to user experience. When looking at for example open positions in the industry many of the job advertisements include user experience either in the job title or at least in the job description. Defining user experience is not anyway always that simple.

3.3.1 Definitions

User experience is defined in different ways in many sources. Sometimes it even feels that there are as many definitions as there are people giving the definitions. When comparing usability and user experience it can be noticed that sometimes they are used almost as synonyms and sometimes they differ from each other. In some cases user experience can be seen as a part of usability (Bevan 2008, 3) and in some cases usability as part of user experience (Morville 2004).

In ISO 9241-2010 standard (2019, 4) user experience is defined as:

> *User's perceptions and responses that result from the use and/or anticipated use of a system, product or service.*

According to the standard user's perceptions and responses include all the emotions, beliefs, preferences, perceptions, comfort, behaviors, and accomplishments that occur before, during, and after use (ISO 9241-210 2019, 4).

Hassenzahl (2008, 12) defines user experience only to be a momentary feeling that takes place while interacting with a product. In the definition user experience is primarily an evaluative feeling between good and bad. When looking at definitions of user experience with a broader scope it can be noticed that Hassenzahl's way of thinking is not that popular. In most cases user experience is seen to include the time spans before, during, and after use.

3.3.2 Time spans of user experience

While most of the user experience takes place during the actual usage, it does not cover all relevant user experience concerns as can be seen in Figure 6. There can be indirect experiences already before the actual usage and after it. These experiences can be formed for example through conversations, recommendations, or advertisements. (Roto et al. 2011, 8.)

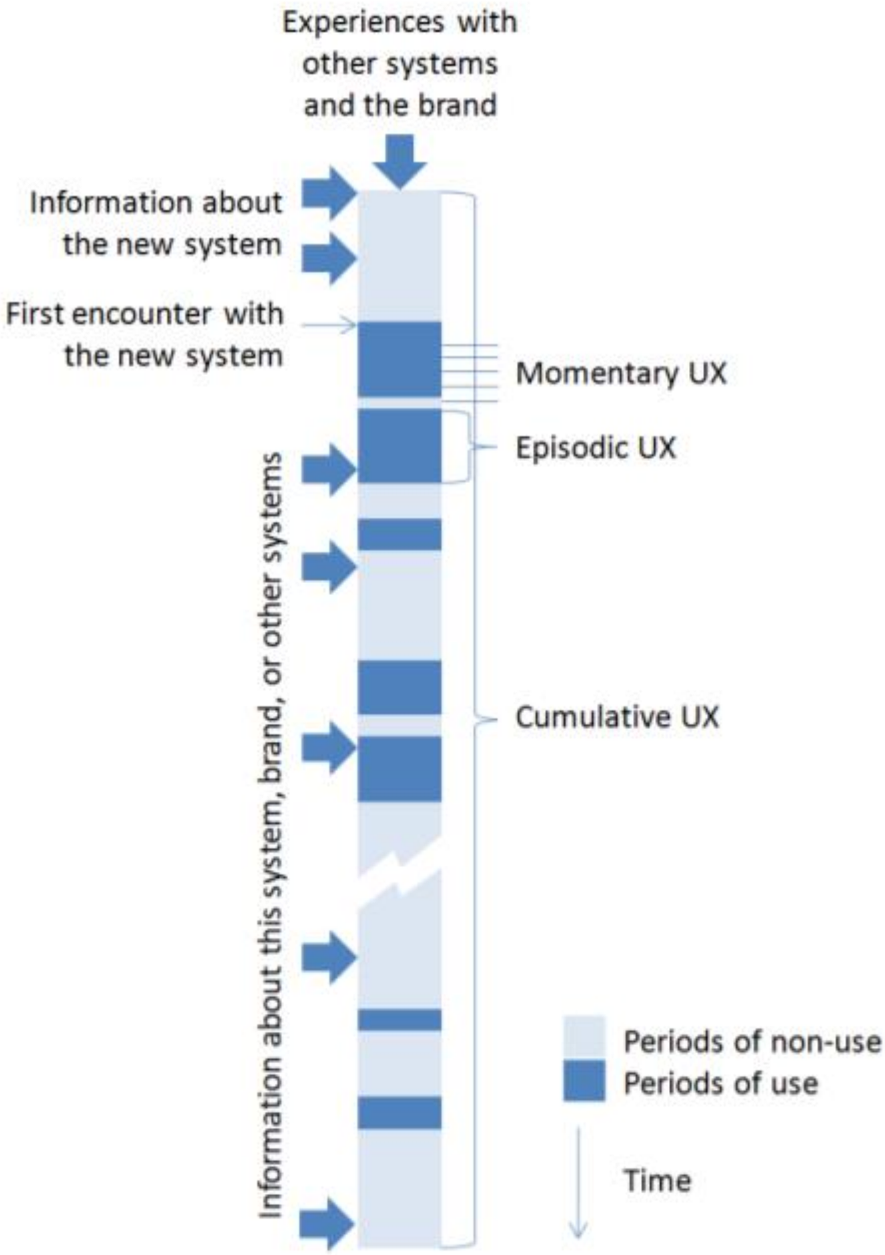


Figure 6: UX over time with periods of use and non-use (Roto et al. 2011, 8).

One can concentrate on different time spans of user experience because of different reasons. Sometimes it is wanted to know what the experiences are during usage and sometimes what they are over time (Roto et al. 2011, 8). Figure 7 shows the time spans of user experience, the terms that describe what kind of user experience is related to them, and the internal process that is taking place in the time spans.

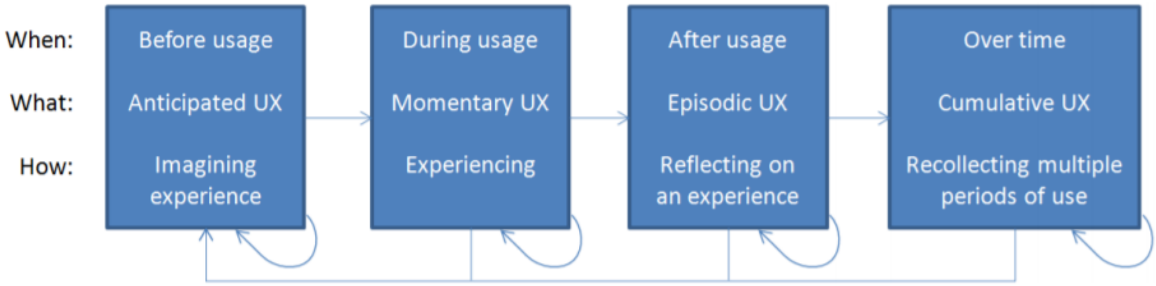


Figure 7: Time spans of user experience (Roto et al. 2011, 8).

UX can refer to a change of feeling during usage (momentary UX), appraisal of a specific usage episode (episodic UX), or views on a product as a whole, after it is used for some time (cumulative UX). Anticipated UX on the other hand may relate to the time before the first use or any of the three other time spans of UX, since a person may imagine a specific moment during interaction, a usage episode, or life after taking a product into use. (Roto et al. 2011, 8.)

3.3.3 UX maturity

Like in any function or practice, all organizations have not adopted user experience design to the same level of maturity (Plewes & Fraser 2015). Figure 8 shows five stages that can be used to compare organizations' UX maturity based on six key indicators.

By addressing the six key indicators it can be evaluated what the organization's level of UX maturity is. First of all there is the timing of UX involvement in the design and development process. The earlier UX is involved, the more mature is the company. Secondly the UX expertise and resources in-house or ability to bring in UX expertise quickly as needed are evaluated. The third indicator is the use of appropriate techniques and deliverables to obtain and understand user input and capture UX design. The leadership and culture in the company are the fourth indicator. It comes to the fact how well the leaders and company as a whole appreciate the value and necessity of UX design from a business perspective. As the fifth indicator the degree to which UX processes are connected and integrated with other corporate processes is evaluated. A degree that enables individuals to work together to create the user experience of the products is preferred. Lastly it is evaluated if design thinking is applied in the broadest perspective possible to drive consistent customer experience. (Plewes & Fraser 2015.)

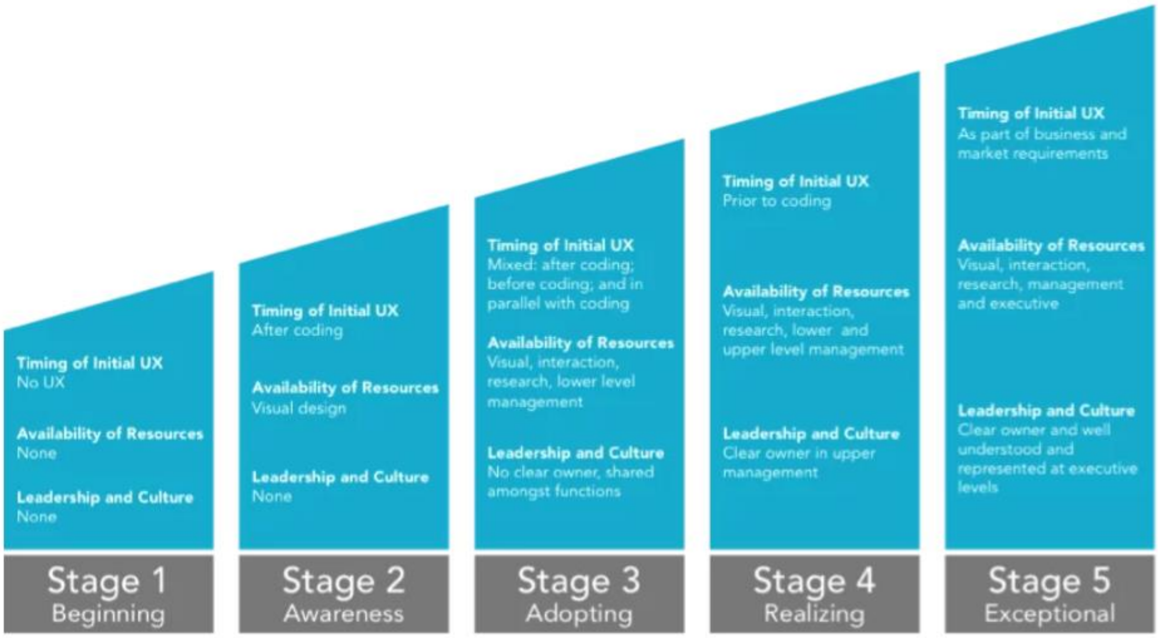


Figure 8: Five stages of UX maturity (Plewes & Fraser 2015).

When thinking about the level of UX maturity at Thermo Fisher Scientific, it can be estimated to be on stage 4, realizing. The timing of UX involvement starts early on in the projects, there is an own in-house UX team, appropriate techniques are used, and UX processes are connected with other design and development processes. Since the world is not yet perfect, there is still some work to do in creating the company wide UX culture, appreciating the value of UX from a business perspective, and applying design thinking to drive consistent customer experience. The level of UX maturity gives anyway good basis for this thesis and for its part this thesis can have an effect on moving the UX maturity closer to stage 5.

3.4 UX metrics

User experience metrics are signals or attributes that show whether the chosen user experience strategy is working or not. With the help of UX metrics it is possible to set targets, track changes over time, and benchmark against iterations of your own product or those of competitors. (Pavliscak 2014.)

UX metrics should not be confused with key performance indicators (KPIs). KPIs reflect the overall business goals such as revenue growth or increased user numbers whereas UX metrics track the user experience (Ratcliff & Kelkar 2019). In some sources UX metrics are also called as UX KPIs but in this thesis the term UX metrics is used.

Tracking UX metrics for products is important and brings benefits. If the UX is not measured, how can it be known how it is performing? How to know if design changes made are for the better or for the worse? (Turner 2018.) There are many rationales to start measuring the UX metrics. According to Meyer (2019) the three most important reasons are stakeholder management, UX benchmarking, and early warning system.

Stakeholder management: UX metrics help to communicate the UX issues and the associated strategic goals more successfully to the relevant decision makers. Without UX metrics, it is basically impossible to prove whether the UX team's actions have been worthwhile and effective. (Meyer 2019.)

UX benchmarking: UX metrics are a powerful business navigation tool that protects from making a wrong turn and wasting valuable resources (time and money). UX data also allows to benchmark the project against internal or external reference data sources (e.g. competitors, if data is available). (Meyer 2019.)

Early warning system: UX metrics reduce the complexity of large amounts of data and give fast and accurate information about the status of UX of the product. With the help of UX metrics it can be quickly seen whether something is wrong and rework is needed. (Meyer 2019.)

With UX metrics the problem of scarcity does not exist. There is an awful lot that can be measured and the number of potential UX metrics is ever growing (Turner 2018). Leech (2011) has listed more than a hundred possibilities what to measure. The idea is anyway not to measure all of the things but to concentrate on the ones that are the most useful in each case.

To help deciding which UX metrics to focus on, they can be categorized to behavioral and attitudinal UX metrics.

3.4.1 Behavioral UX metrics

Behavioral UX metrics are related to what the users are doing. They express in numbers what users are effectively doing and how they interact with the product. (Meyer 2019.) In today's user research world it is critical to understand what users are doing and how they are using the products (Ratcliff & Kelkar 2019).

Nowadays the data related to behavioral UX metrics can quite often be collected fully automatically without the intervention of an observer or an interviewer. Therefore behavioral UX metrics are fairly simple and inexpensive way to start collecting UX metrics data. (Meyer 2019.)

In all cases it is anyway not possible to collect the data automatically. This is the case most of the times also with laboratory automation devices. In these cases, task based usability testing is a standard method to gather the information (Ratcliff & Kelkar 2019). Users are given a certain set of tasks and based on decided UX metrics it is observed how the users perform the given tasks.

The following metrics are examples of behavioral UX metrics (part of them will be examined closer later in this thesis): abandonment rate, page views, problems and frustrations, task success rate, time-on-task, search vs navigation, and user error rate (Meyer 2019; Ratcliff & Kelkar 2019).

3.4.2 Attitudinal UX metrics

Attitudinal UX metrics are related to what the users are saying and feeling. The attitudinal UX metrics measure how users feel or what they say before, during, or after purchasing a product. (Meyer 2019.) They are also related to how users feel, what they say before, during, or after using the product, and how this all is affecting their brand perception (Ratcliff & Kelkar 2019).

Attitudinal UX metrics are where qualitative data, such as appearance, loyalty, trust, and usability, is quantified. Quantifying the opinions of users can be challenging but there many different methods that assign a numerical score to attitudinal data. (Ratcliff & Kelkar 2019.)

The following metrics are examples of attitudinal UX metrics (part of them will be examined closer later in this thesis): loyalty, usability (or ease of use), credibility, appearance, system usability scale, net promoter score, and customer satisfaction (Meyer 2019; Ratcliff & Kelkar 2019).

3.5 ROI of UX activities

Return on Investment (ROI) is evaluating how efficient an investment is (Chen 2020). It is a performance measure of what is got back compared to what is put in (Duermyer 2020). When calculating ROI, the return of the investment is divided by the cost of the investment. The result of this calculation is presented as a ratio or a percentage. (Chen 2020.)

Designing user experience that fits customer’s expectations and at the same time meets business goals involves effort and investment but it is definitely worth it. UX design has proven to have a positive return on investment. Often UX is the factor that differentiates successful products from the unsuccessful ones. (Tej 2019.)

According to different studies the positive ROI for UX design has been measured to produce significant savings. There are studies that show that one dollar invested in UX brings 100 dollars in return, 90% reduction in support costs can be achieved after usability testing, 70% of projects fail because of lack of user acceptance, and that 50% of developers' time is spent on re-work that could be avoided. These all have led to measurements where even 93% of executives have thought that improving user experience is a top strategic priority. (Tej 2019.) Especially interesting for laboratory automation product development is the amount of re-work that could be avoided since those savings can be directly seen inside the product development department. If re-work can be avoided the resources can be directed to other projects to increase the efficiency of work.

Sometimes the value of UX design might be questioned since all design work does not have value and also all valuable design work is not of equal value. In UX design the ROI is often about eliminating poor design. It might not be about how much more income to generate but more about how to reduce the costs. (Spool 2018.) With UX design certain frustration costs can be reduced. These costs include costs due to lost sales revenue, increased support costs, lost productivity costs, wasted development re-work, and unused feature development. (Spool 2017.)

The ROI of UX activities should be considered more as an estimation than sophisticated financial forecasts. They are estimates of how much value the company is getting as a result of the investment. It is important to understand that the ROI of UX is not always just all about money. It is more about demonstrating that design that improves the user experience has a positive impact on the business goals which can sometimes be also not directly related to money. (Moran 2020.)

Since calculating the ROI of UX activities can be challenging it can be categorized into soft and hard dollars. The soft side includes results such as increased customer loyalty, increased word-of-mouth referrals, increased efficiency, decreased user errors, and productivity gains. On the hard side there are increased sales, cost savings from support and training, less money spent on development and redo cases etc. (Philips 2018.)

4 RESEARCH METHODOLOGIES

The fourth chapter describes the research methodologies that are used in the empirical research phase and some background behind them. The chapter introduces Lean, Six Sigma, and Lean Six Sigma methodologies, and explains how to utilize lean UX canvas, DMAIC process, interviews, and usability testing in the research and development.

4.1 Lean

Lean process is a method for eliminating or limiting waste. The goal is to increase efficiency and productivity without consuming more resources. Benefits of lean include e.g. completing activities faster without performing poorly, getting things done with less human effort, and consuming less material without sacrificing the quality. (Bhanushali 2019.)

Instead of being a quick fix lean is rather a cultural transformation that changes the way of working. Lean methodology recognizes two important factors: value added activities and non-value added activities. Value added activities are the ones that customers care for and they boost the value of organization's offerings. In lean methodology the aim is to concentrate on the value added activities and remove the non-value added activities. (Bhanushali 2019.)

There are five principles in the lean methodology. The principles can be seen in Figure 9.

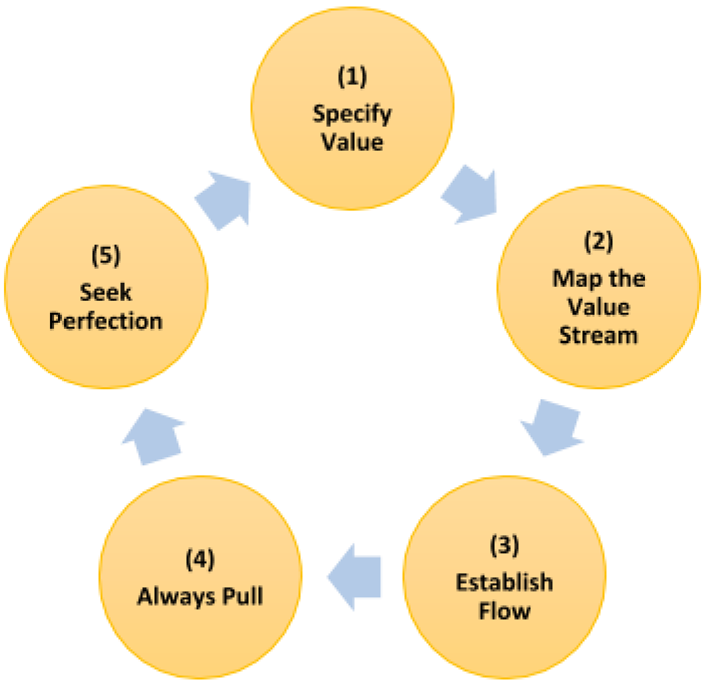


Figure 9: Principles of lean methodology (Bhanushali 2019).

1. Specify value

Defining value is the first step in lean thinking. The value is specified and quantified from the user's perspective. Users might be customers or employees but they are the ones defining what value means. (Bhanushali 2019.)

2. Map the value stream

The second principle focuses on value stream mapping. The value stream is the series of activities and the process in use. Both value added and non-value added activities are included in the value stream. When applying lean methodology the bottlenecks and waste can be eliminated from the value stream. (Bhanushali 2019.)

3. Establish flow

The third principle is to improve the workflow and enhance value generating activities. Enhanced workflow will help in moving products, services, and information down the value stream without delays. The workflow is enhanced by eliminating stops and delays while improving the flexibility and reliability of the process. (Bhanushali 2019.)

4. Always pull

In the fourth principle the customer must pull value from the next upstream activity once the flow generation is started. Each pull is initiated by customer demand which ensures that the processes follow an unbreakable sequence. (Bhanushali 2019.)

5. Seek perfection

According to the fifth principle there is a need to constantly seek improvement and to perfect the system. Business operations and processes are iterative and under constant improvement. In an ideal world a state, where best value possible with zero waste in the process is delivered, is reached. (Bhanushali 2019.)

4.2 Lean UX canvas

The lean UX canvas (Figure 10) is created to help teams frame their work as a business problem to solve rather than a solution to implement. The business problem is dissected into its core assumptions that are weaved into hypotheses. Finally, from the hypotheses experiments are designed to test the riskiest hypotheses. (Gothelf 2016.) The lean UX canvas is derived from the business model canvas created by Osterwalder and Pigneur (2010, 44).

In the lean UX canvas business problem, business outcomes, users, and user outcomes and benefits (boxes 1-4 in Figure 10) describe the current situation and outcomes. Solutions (box 5) tries to answer how the outcomes can be reached. Hypotheses, what's the most important thing we need to learn first, and what's the least amount of work we need to do to learn the next most important thing (boxes 6-8) try to answer the question how to find out if the solutions were right. (Gothelf 2019.)

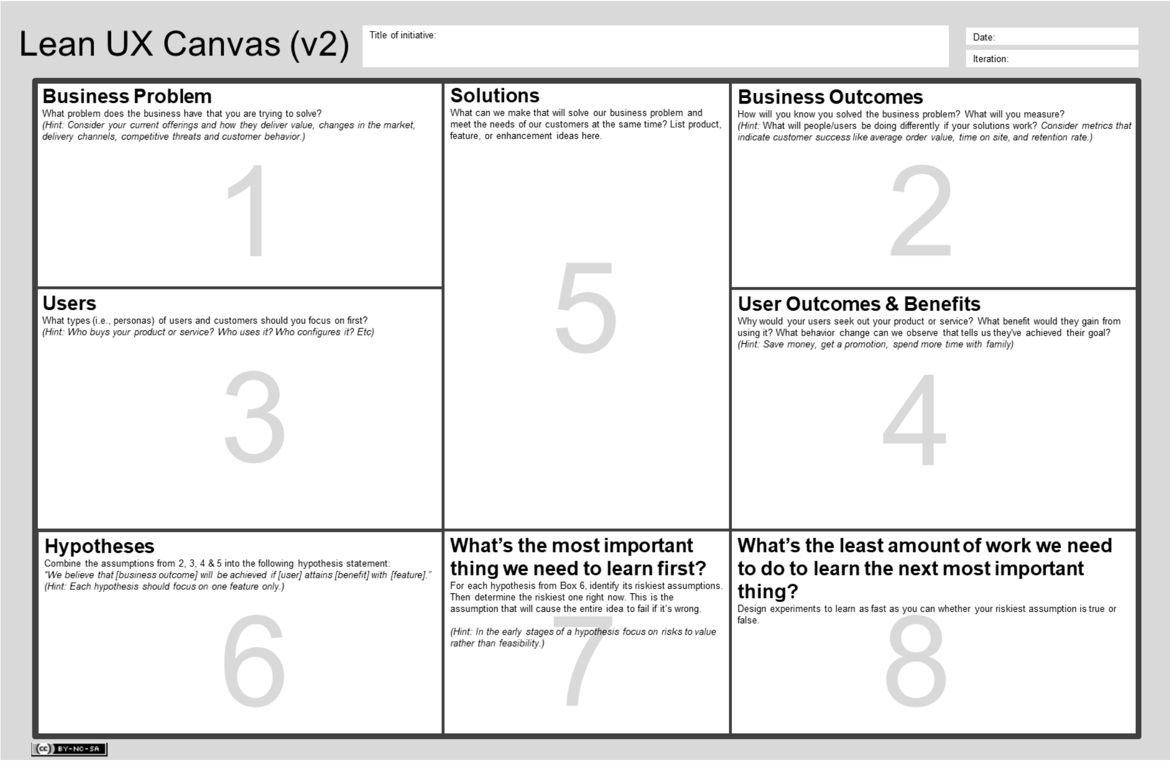


Figure 10: Lean UX canvas (Gothelf 2019).

Gothelf (2019) explains that the lean UX canvas serves for several purposes. Lean UX canvas is a facilitation tool for cross functional teams designed to create a customer centric conversation about the work the team is doing. Its goal is to help the team focus on why they are doing the current work. Lean UX canvas also provides an initial recipe for teams to follow in the early stages of agile adoption. It is a kind of insurance policy to ensure that learning takes place in every sprint. It exposes the gaps in the team's understanding of the problem they are solving, who they are solving it for, and why they believe their solutions will work. Lean UX canvas acts also as the first step in the shift of the conversation from outputs to outcomes. (Gothelf 2019.)

4.3 Six Sigma

The basic formula of Six Sigma is simple: Give the customers what they want, when they want it, at a competitive price, and your company can become the provider that produces highest quality products with lowest cost (Brue 2015, 221).

Six Sigma is a methodology where a set of techniques and tools is used to improve product or service quality by identifying and reducing (or eliminating completely) the causes of defects or errors and minimizing variability in processes (Brue 2015, 1). Six Sigma methodology can be used in all business processes and it can generate significant financial results (Brue 2015, 13).

Customers are the starting point in Six Sigma. They essentially define the quality for products and services. It is good business to focus on meeting customers' expectations since failing to meet them at all might end up losing the business completely. (Brue 2015, 16.) In Six Sigma the extent to which the products and services meet customer expectations is measured. The following questions help in that (Brue 2015, 15-16):

- What is it that the customers want?
- What do the customers not want?
- What is done to deliver products that meet customers' critical-to-quality expectations?
    - What processes are involved in meeting the critical-to-quality expectations?
    - What is happening in those processes?
- How well are the customers' expectations met?

When getting started with Six Sigma there are some principles that should be taken into consideration and a few that should be avoided. The principles that are recommended to be taken into consideration include e.g. focusing on results, embracing customers, planning for success, and establishing project baselines and goals. In terms of decreasing costs and increasing profits it is important to know both the current situation and the desired outcome. Customers must be paid attention since meeting customer expectations is an essential factor leading to business growth. With right type of planning and knowing the baselines and goals, the progress of the projects can be measured and goals reached easier. Communicating the project commitment throughout the company and demonstrating the commitment of company leaders help showing all employees that also company leaders fully support reaching the goals. Empowering the key human resources and providing on-site mentoring for them enables the right people to be picked to lead the project teams and allows them to have full control to their work in a way that they know support is anyway available when needed. For the selection of right people and projects help from outside implementation partner can be used. Running Six Sigma projects needs patience since the projects require front end commitment of time, resources, and training in order to deliver results. Claiming and advertising early wins is encouraged since celebrating each milestone of success will help keeping the enthusiasm of the team high. When thinking about the results and wins benchmarking helps conducting the right cap analysis to know the current position. Getting advance buy-in from controller is also recommended since agreeing with controller how to calculate real savings is essential to make a difference between tangible and intangible savings. (Brue 2015, 46-48.)

By following these above mentioned recommended principles there is a high chance for the Six Sigma projects to be successful. The chances become even higher when one remembers what to avoid. Six Sigma should not be made a massive training exercise. Rather than concentrating on constant training it is more important to learn how to implement and get started with the implementation. It is also not feasible to train all of the employees right away. Instead of training everyone it is more effective to start with the right people and right projects. Resources should not be focused on reworking training material either but to get busy implementing. Letting controller complain about the savings calculations can be avoided when project leader and controller are in agreement right from the beginning on how to define the savings in the projects. In Six Sigma techniques all the steps are in for a reason. The steps should not be skipped. Instead they should be used to provide the information needed to solve the problem in question. (Brue 2015, 48-49.)

4.4 Lean Six Sigma

Lean Six Sigma combines the strategies of Lean and Six Sigma. Lean principles help to eliminate or reduce waste in processes while Six Sigma focuses on reducing variation in processes. Together in Lean Six Sigma the principles help to improve the efficiency and quality of the process. (Rastogi 2018.)

Under the tenets of Lean Six Sigma any use of resources that is not creating value for the customer is considered waste and should be eliminated. The combination of lean management concepts and Six Sigma tools and techniques reveal what processes are prone to variation and then reduce those variations to ensure continuous improvement. (Kenton 2018.)

Lean Six Sigma uses belt levels to denote expertise in using the methodology. The belt levels generally start from yellow belt and end to master black belt. Yellow belt means that the person has awareness of Lean Six Sigma. Green belt is a project member who is focusing on the use of tools and application of DMAIC and lean principles. Black belt is a full time project leader of Lean Six Sigma projects. Master black belt means a black belt with at least two years of experience. The persons having master black belt are also able to teach Lean Six Sigma to others. (Kenton 2018.)

4.5 DMAIC process

In general there are two forms of Lean Six Sigma used: DMAIC (define, measure, analyze, improve, control) and DMADV (define, measure, analyze, design, verify). Here DMAIC process is introduced. It consists of five phases (Brue 2015, 8):

1. Define: In define phase the project goals and customer (both internal and external) deliverables are determined.
2. Measure: In measure phase the product characteristics are identified, the process is mapped, measurement systems are evaluated, and the current performance of the process is estimated.
3. Analyze: In analyze phase the variables are evaluated and reduced through analysis and hypothesis testing. Also vital few factors for process improvement are identified.
4. Improve: In improve phase the variable relationships among the vital few factors are discovered, operating tolerances are established, and measurements are validated.
5. Control: In control phase the ability to control the vital few factors is determined and process control systems are implemented.

4.5.1 Define

The purpose of the define phase is to determine the purpose, scope, and objectives of the project, to collect information on customers and current process, and to specify the deliverables that are produced during the project. The design phase starts by addressing the following questions (Brue 2015, 53):

- What is the problem that needs to be focused on?
- Who are the customers?
- What are the critical factors to the customers?
- What are the processes involved?

- What are the critical factors to the processes?
- What is the goal?
- What is the timeline for achieving the goal?

The phase includes several activities. In addition to defining the problem scope and planning the project (resources, team, etc.) the most important areas to focus are the voice of customer, critical-to-quality requirements, and project metrics (Brue 2015, 54).

The input of the customers is known as the voice of the customer (VOC) (Brue 2015, 17). Finding out the VOC can happen in multiple ways. Requirements and specifications can be provided by business customers. Sales representatives may report on what customers want or need, or what does not matter to them at all. Customers can also be asked directly about their needs and wants, and what they think about the products in market. Another way to learn about VOC is to stay informed what competitors are doing with their products. (Brue 2015, 18.)

Critical-to-quality (CTQ) requirements are the requirements that are the most important to the customers. When products or services are wanted to be improved it is important to identify the factors that customers consider critical to quality since the CTQ requirements form the definitions of quality for the products or services. (Brue 2015, 18.)

When CTQ requirements are known project metrics should be defined to measure how well meeting the CTQ expectations is going. The importance of metrics is that when processes can be measured, they can also be understood. When processes are understood, they can be corrected, controlled, and improved. (Brue 2015, 21.)

4.5.2 Measure

In the measure phase the purpose is not to solve the problem but to understand exactly what the situation is, what is known, and what is not known. It is the beginning of the problem solving journey that starts by exploring the output of the problem and by listing potential causes for the problem. (Brue 2015, 77.) Critical tools that can be used in the measure phase include e.g. process mapping, output of the problem – potential cause of the problem –matrix, failure modes and effects analysis, measurement system analysis, and process capability analysis (Brue 2015, 114).

In order to better understand the processes and to find out causes for the problem value stream mapping and SIPOC (suppliers, inputs, process, outputs, customers) (Figure 11) can be used.

Value stream mapping (VSM) is a tool to improve business efficiency. It is known as material and information flow mapping to analyze and evaluate certain work processes. VSM is primarily used to identify, demonstrate, and decrease waste but also to create flow in the manufacturing process. (Verma & Sharma 2020, 40-41.)

Figure 11: SIPOC diagram (Brue 2015, 58).

SIPOC diagram is a tool that is used to identify all relevant elements of a process improvement project. It starts as a row of five headings but as the project progresses lists and other details are added on it. (Brue 2015, 20.) The five components of SIPOC are suppliers, inputs, process, outputs, and customers. Suppliers are the sources of inputs into the process. Inputs consist of material, information, and other resources. Process describes the activities that are grouped to transform inputs into outputs. Outputs represent the products or services that are resulting from the process. Customers are the recipients of the outputs. (Brue 2015, 19.)

4.5.3 Analyze

The two purposes of the analyze phase are to evaluate and reduce the variables by using analysis and hypothesis testing, and to identify the vital few factors in order to identify root causes for the problem. The analyze phase helps to generate ideas for improvement. (Brue 2015, 115.)

The major transition in analyze phase is to start using data to know rather than relying on thinking, believing, and feeling. Whatever tools that help sorting through the potential causes of the problem and linking them to the vital few factors can be used. (Brue 2015, 146.) Possible tools to be used include e.g. cause and effect diagrams. Five whys is one cause and effect technique used for root cause analysis. In short, a problem exists and the question why is asked five (or as many as needed) times in order to find the root causes behind the problem. When the root causes are known they can be fixed and it is possible to totally remove the original problem or at least decrease the effects of the problem. (Leijala 2015, 77.)

At the end of the analyze phase root causes of variation and the significant factors to be further investigated should be identified and verified. (Brue 2015, 146.)

4.5.4 Improve

The improve phase is initiated by selecting the product or process performance characteristics that must be improved in order to achieve the goal. To improve the process one must be ready to develop, test, and implement multiple solutions. Then with data it must be demonstrated that these solutions work. First ideas are generated to improve the process, then the ideas are analyzed and evaluated, best potential solutions are selected and tested, the solutions are planned and implemented, and finally the results are validated with data. (Brue 2015, 147-148.)

Design of Experiments (DOE) is the main tool in improve phase (Brue 2015, 177). It is a structured way to conduct controlled tests and analyze the test results. Experiments can be used to discover or study an effect of a variable, or to test or develop a hypothesis. There are several ways to conduct experiments to test the effects of the variables. Traditional approach is to test one variable at a time but instead of running a series of experiments also all factors can be tested at the same time. (Brue 2015, 153-154.)

DOE consists of three pieces: planning the experiment, conducting the experiment, and analyzing the experiment results. Planning is the most important piece since it defines the logic of the experiment and what is wanted to be achieved with it (Brue 2015, 155). Conducting the experiment is simply a matter of following the plan although it can be challenging and time consuming (Brue 2015, 166). Analyzing the experiment results enables to identify the most influential factors associated with critical-to-quality characteristics, to understand interactions between various factors, and to get information to quickly improve the process (Brue 2015, 177).

4.5.5 Control

The purpose of control phase is to control the potential causes of the problem to ensure a sustained output of the problem. The goal is to use process data to monitor the process and adjust it as necessary to make it perform as intended. The control phase is usually initiated by selecting the performance characteristics that must be improved to achieve the goal. (Brue 2015, 208.) Basic steps of control phase include checking the results of the improvement, building a control plan, implementing the control plan, transition to process owners, and closing the project (Brue 2015, 180).

The control plan is a management tool that ensures that the improvements do not fade away after the improved process is transferred back to the process owners (Brue 2015, 181). The process owners must be communicated the documented process and its improvements and benefits so that they understand how the process has been changed and what it means in terms of responsibilities (Brue 2015, 199).

4.6 A3 report

The A3 report is a problem solving and continuous improvement tool which is named after the A3 size sheet of paper. The reason for the naming is the intention to fit the whole report in a single page. (Thakur 2019.)

The A3 report is especially useful for completing less complex or short Lean Six Sigma projects. It might not be that useful when a larger improvement project or high complex project is undertaken in a DMAIC environment (Thakur 2019) but it can be anyway used to give a summary of DMAIC process. DMAIC model can also be used to organize the A3 report to make it easier to follow the logical flow of information.

Typically an A3 report includes the topics of problem background, current situation, analysis, goal, recommendations, implementation plan, follow up, and results report. Background section is a short description of the problem highlighting the measures that are used. Current situation describes the current state of the problem visually. Analysis is used to determine the root causes of the problem. Goal depicts visually the desired results without the problem occurring. Recommendations describe the solutions that need to be implemented. Implementation plan lists the tasks that are needed to be performed, their starting times, durations, responsibilities, and statuses of completion of the tasks. Follow up of the tasks ensures that the benefits of the solutions are maintained. Result report presents a charted progress with implementation and measures. (Thakur 2019.)

These topics can be presented as separate sections in the A3 report, or as in the case where DMAIC organizes the report, the topics can be included under the reporting of define, measure, analyze, improve, and control phases.

4.7 Interviews

Interviews have become probably the most used method for collecting information. This is not a surprise since it is natural to ask when something is not known beforehand. (Ruusuvuori & Tiittula 2005, 7.)

Traditionally research interviews have been divided into structured and unstructured interviews. Sometimes they are also called standardized and non-directive interviews. In structured interviews the questions are presented every time exactly the same way in the same order. Structured interviews are commonly used e.g. in surveys. Unstructured interviews are more like discussions since they do not have any prearranged questions. Both the interviewer and the interviewee can guide the discussion and raise topics of their interest. (Ruusuvuori & Tiittula 2005, 10.)

In between structured and unstructured interviews there are semi-structured interviews. In semi-structured interviews it is natural to go through the same topics in every interview but the wording and the order of the questions may change. Often questions are more open ended than in structured interview. (Ruusuvuori & Tiittula 2005, 10.)

Interviews can be held as one-on-one interviews or as group interviews. One-on-one interview is carried out between an interviewer and an interviewee. It is purely qualitative method which allows to get meaningful data when the right questions are asked. Group interview allows a number of participants to be interviewed at once. Group members may influence each other with their comments which can be both a good and a bad thing. (Bolderston 2012, 68.)

4.8 Usability evaluation

There are various methods for evaluating products during the design and development process. Different methods can be used for prior and during the product development, and after the product is already published (Faulkner 2000, 137).

Faulkner (2000, 137) states that in order to produce a good product it is not enough to ensure that the product and the human using it both perform as efficiently as possible. The use of the product must also be satisfying for the user. One way to help achieving this is to evaluate the product throughout the development process and this way set the direction for design and redesign activities. (Faulkner 2000, 137.)

When performing the usability evaluation it is important that there is also a clear set of project goals. Without them one can get easily stuck in the evaluation results and start perfecting features that are actually not that important what comes to the end product. Usability evaluation should anyway not be overlooked. It is iterative and should take place as each design stage occurs. (Faulkner 2000, 138.)

The following sub chapters introduce two types of usability evaluation (formative and summative) that are relevant in the laboratory automation product development.

4.8.1 Formative usability evaluation

Formative usability evaluation is user interface evaluation that is conducted iteratively throughout the design and development process. The main target of formative usability evaluation is to explore the strengths, weaknesses, and use errors of the user interface design. (IEC 62366-1 2015, 9.)

The purpose of formative usability evaluation is to help the design process and it should be started at a very early stage of the design and development process, since possible product changes are cheaper the sooner they are taken onboard. (Faulkner 2000, 138.)

Formative usability evaluation involves working closely with users and gathering feedback about their opinions of the evaluated product. The nature of these opinions can depend based on what evaluation method is being used. (Faulkner 2000, 138.) Methods for formative evaluation can consist e.g. of various types of reviews, heuristic analysis, cognitive walkthrough, and usability testing. The data received on the other hand can include e.g. customer preference survey responses, user comments, and usability test participants' ratings and rankings. (IEC TR 62366-2 2016, 53-54.)

Formative usability evaluation that includes usability tests does not typically have formal acceptance criteria. The purpose is to iterate the design and to achieve good enough quality level to increase the likelihood that the summative usability evaluation can be conducted successfully. (IEC 62366-1 2015, 35.)

4.8.2 Summative usability evaluation

Summative usability evaluation is performed at the end of the product development. It is the last evaluation performed for a product and its goal is to obtain objective evidence that the use of the device can be performed safely without use errors. (IEC 62366-1 2015, 10.) It gives a measurement of what the improvements in performance of the system actually are (Faulkner 2000, 139).

The method used in summative usability evaluation is mainly usability testing that is performed with large enough amounts of participants from the intended user groups of the product (IEC TR 62366-2 2016, 57).

Summative usability evaluation is part of activities involved in verifying and validating the overall medical device design and for this reason it is performed for a complete design (IEC 62366-1 2015, 12). It can anyway sometimes be performed for one completed section of the product at a time so it might not always be the entire product with all of its consumables (e.g. separately for information for safety). (Faulkner 2000, 139.)

5 RESEARCH AND ANALYSIS

The research and analysis phase of this thesis was iterative and consisted of two main rounds that are described in this chapter. Also inside the two main rounds iterative principles were followed since the research actions were built on top of the previous findings. In the development the principles of lean, Six Sigma, and Lean Six Sigma were taken into account.

The performed data collection, research, and analysis activities are reported in the following chapters. The chapters first give an overview of the data collection, describe the use of lean UX canvas and collection of preliminary list of UX metrics, and then continue explaining the use of DMAIC process by taking a closer look on the problem space and current situation, analyzing the root causes of the problem, designing the experiments performed (planning the experiments, conducting the experiments, and analyzing the results of the experiments), and proposing recommended actions for follow up.

5.1 Data collection

The data used in the research and development of this thesis was mainly collected during a four-month-long period of time. This time frame contained two rounds of research and data collection. The research methods used were qualitative since quantitative methods were not seen appropriate for the topic in question.

In the first round the first set of data was created by using the lean UX canvas and by collecting a preliminary list of UX metrics from other fields of industries since lists of UX metrics directly for medical devices or especially for laboratory automation devices did not exist publicly. In the lean UX canvas existing findings of user research already performed during the past few years were utilized. For the creation of preliminary list of UX metrics a comprehensive set of publicly available UX metrics was reviewed and based on certain requirements a presumably appropriate set of UX metrics was selected to be researched further on the second round.

In the second round DMAIC process was used. The initial presumptions from the first set of data were taken as a basis for the second round. During the second round as part of DMAIC process data was collected in various experiments. The data collection methods included end user interviews, product management interviews, usability testing, and going through lists of tender requirements. The preliminary list of UX metrics was put under evaluation, analyzed based on interview and usability evaluation results, and modified further. At the same time it was analyzed if monetary value can be given to usability engineering and user experience design activities.

In all the interviews and usability evaluations notes were taken and in the usability evaluations the behavior of the test participants was observed. Summaries of these notes can be seen in the appendices of this thesis. In the following chapters the performed data collection, research, and analysis activities are given a closer look.

5.2 First round of research execution – Initial presumptions

The first round of research consisted of addressing the business problem through lean UX canvas and of collecting a preliminary list of UX metrics.

5.2.1 Lean UX canvas

A lean UX canvas was used to address the business problem, to find solutions to the problem, and to see towards which direction to move next. It acted as a good starting point for the research when the first presumptions and information could be collected into one canvas. Filled in lean UX canvas is seen in Figure 12 and a larger size version of the canvas can be found from Appendix 1.

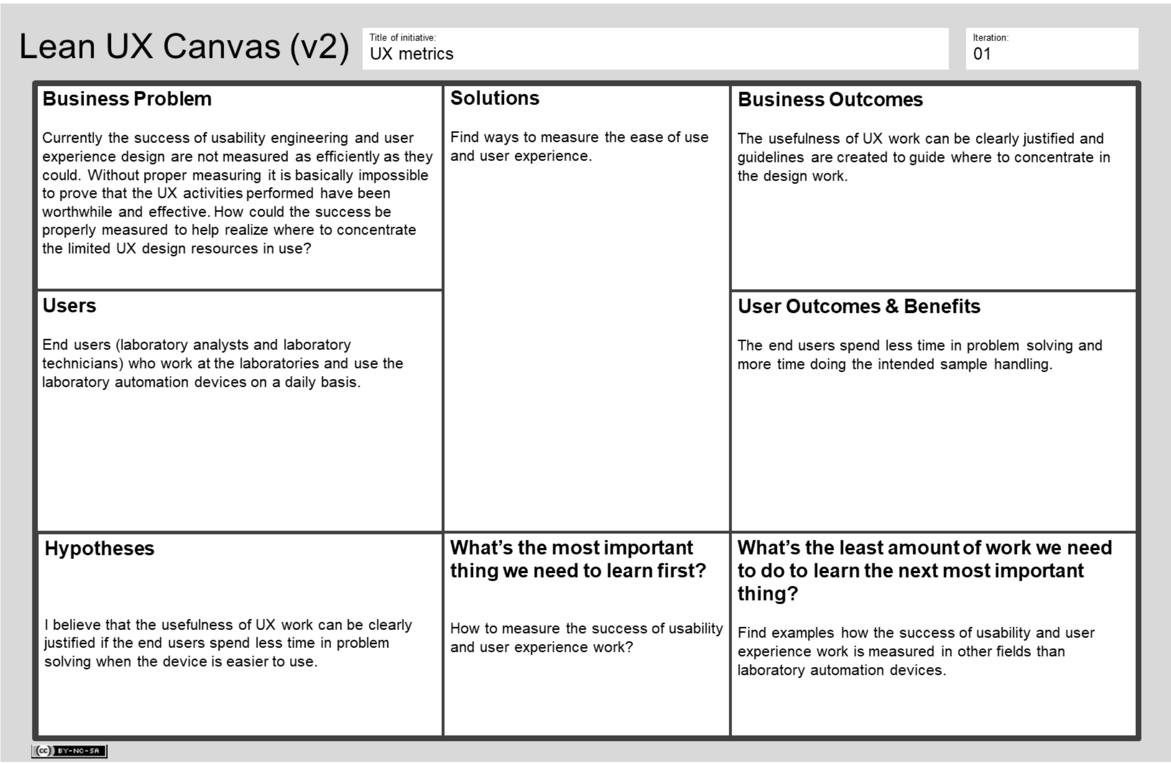


Figure 12: Lean UX canvas - UX metrics.

The business problem statement is related to measuring the success of usability engineering and user experience design. It is currently not measured as effectively as it could be and without proper measuring it is basically impossible to prove that the performed UX activities have been worthwhile and effective. It all comes to the question, how could the success be properly measured to help realize where to concentrate the limited UX design resources in use? As an outcome to the business problem the usefulness of the UX work should be clearly justified and guidelines to guide where to concentrate in the design work should be created.

The users in this case are the end users who work at the laboratories and use the laboratory automation devices on a daily basis. They are mainly laboratory analysts and laboratory technicians. The benefit for them after solving the business problem is that they would not need to spend that much time for problem solving anymore and they could concentrate on the intended sample handling instead.

The solution for solving the problem and to reach the business and user outcomes is to find suitable ways to measure the ease of use and user experience, which also forms the hypothesis, that the usefulness of UX work can be clearly justified if the end users spend less time in problem solving when the device is easier to use. When the laboratory automation devices are easier to use there will not be as many problems occurring in the first place.

In order to reach the solution and to prove the hypothesis to be correct it is important to learn how to measure the success of usability engineering and user experience design. The first step towards this is to find examples of suitable metrics that can be used to measure the success of usability and user experience work. Since finding those examples from the field of laboratory automation was impossible, other fields of industries were explored. There seems not to be UX metrics which are proven to work with laboratory automation devices publicly available.

5.2.2 Preliminary list of UX metrics

In order to create a metrics driven approach to evaluate the success of usability engineering and UX design there are five steps that need to be followed (Nguyen 2020).

1. Understand business goals.

There are loads of possibilities what metrics to track but the right metrics to choose are the ones that impact the company's business goals (Zibell 2017). The following questions help in choosing the right metrics (Nguyen 2020):

- What are the most important metrics for the business or for the product?
- What are the business values of the new product or feature?
- How is the success of the new product or feature measured?

Some examples of company's business goals which can be also recognized in laboratory automation product development are e.g. growth, profitability and efficiency, and customer service and retention. Growth means increasing revenue through new markets, new products, or acquisitions. Profitability and efficiency are linked to doing things faster, better, or cheaper in order to increase profit. Customer service and retention on the other hand are related to the assumption that satisfied customers are loyal and valuable over time. (Zibell 2017.)

2. Translate business goals to UX metrics.

It is hard to attribute the improvement in business goals to solely UX design as there are multiple factors affecting them. For this reason UX metrics, that can be influenced, must be defined. An example for a UX metric that is connected e.g. to retention could be customer satisfaction rating. (Nguyen 2020.)

3. Set targets.

When the metrics are defined, it needs to be defined what kind of changes to these metrics will be considered a success. The targets are adjustable but they force to think what success looks like in advance. The targets also hold designers accountable to take actions if the results are far away from the expectations. (Nguyen 2020.)

4. Measure baseline.

Measuring a baseline gives a starting point where to compare the future measurements. Some methodologies to collect baseline metrics include e.g. surveys, observations, interviews, and analytics (Nguyen 2020).

5. Measure after design.

After making changes in the design the metrics are measured again. If results fall below targets, they should be analyzed to see if design tweaks to improve the results are needed. If the results hit the targets the changes, new feature designed, or the product can be held successful. (Nguyen 2020.)

Out of these five steps understanding the business goals and translating them to UX metrics are fairly stable for a product with an established business model. Also setting the targets should be quite straightforward and not require a lot of time. For the baseline and after design measurements periodic assessment (e.g. usability test or survey) of the product should give the needed results. (Nguyen 2020.)

This thesis focused on the first two steps in order to create a suitable list of metrics for the design and development of laboratory automation devices. The three latter steps enter the picture in a later phase of product development but were not part of this thesis. There is no universal truth what metrics are the best. Instead it depends of the business goals and the industry in question.

Creating a preliminary list of UX metrics started with conversations inside the UX team. The aim was to discuss about UX metrics that cover both behavioral and attitudinal factors, work in laboratory automation product development, and can be tested in the later phases of this thesis. Use safety, use effectiveness, ease of use, problems that users encounter, previous experiences with laboratory automation devices, and possible recommendations from colleagues emerged from the discussions. It was clear that these factors should be covered in the metrics in some form, when a preliminary list of UX metrics was collected.

After going through several possibilities the seven most important UX metrics from Meyer (2019) were selected as the preliminary list of UX metrics. They acted as a baseline and were given a closer look later in the research to make sure if they really work as metrics in the design and development of laboratory automation devices. The seven UX metrics were chosen because they work for any type of projects (containing both software and hardware), they can be connected to many common business goals, they have an impact on many design decisions, they include both behavioral and attitudinal metrics, they cover the factors that were found important in the discussions inside the UX team, and they were also cited in other sources. These seven preliminary UX metrics consist of four behavioral UX metrics (task success rate, time on task, search vs. navigation, and user error rate) and three attitudinal UX metrics (system usability scale (SUS), net promoter score (NPS), and customer satisfaction (CSAT)). Here below the selected UX metrics are introduced in a more detail.

Task success rate

Task success rate is a very frequently used metric. It can be used whenever a task has a clearly defined endpoint. Task success rate measures the number of correctly executed tasks compared to the total number of attempts. It does not give information why a user could not successfully complete a task but it is anyway a very valuable indicator. (Meyer 2019.)

The task success rate is presented as a percentage. If for example nine out of ten users are able to complete a task, the task success rate is calculated in the following way:

9 / 10 = 90 %

Task success rate can be measured both for first time use and when the user already has more experience with the product. Comparing these two allows to see if there are changes in the metric. The higher the task success rate is, the better the user experience. (Meyer 2019.)

Time on task

Time on task describes the time that users need to complete a task successfully. Usually the average time on task among all users is communicated as the final UX metric. Basically the shorter the time on task is, the better the user experience. (Meyer 2019.)

Search vs. Navigation

When using graphical user interfaces users typically have two options to reach their destination. They can reach their destination either via navigation or by using the search function. In many cases, the less the search function is used, the better the user experience. This can anyway vary case by case. (Meyer 2019.)

The metric is presented as percentage ratio of how many times out of completed tasks either navigation or search function was used to reach the desired outcome. (Meyer 2019.)

User error rate

User error rate describes the number of times a user made an error while working on a task. It gives an idea of how clear and user friendly the product is. The higher the user error rate score is, the higher is the number of usability problems. (Meyer 2019.)

There are two frequently used ways to calculate the user error rate. Error occurrence rate is used if a task only allows one potential error or if there are several but only one of them is wanted to be measured. Error occurrence rate is calculated as the total number of errors that occurred for all users divided by the total number of possible errors for all users. E.g. if five out of fifty users made an error the error occurrence rate is 5 / 50 = 10 %. (Meyer 2019.)

Another way to calculate the user error rate is to use the error rate. It can be used if multiple errors per tasks are possible or if multiple errors are wanted to be measured. In error rate the number of errors is divided by the total number of attempts. (Meyer 2019.)

System usability scale (SUS)

System usability scale is kind of a quick and dirty tool for testing the usability of a product. It can be used on any kind of product and it has become an industry standard with references in hundreds of publications. (Sauro 2011.)

SUS consists of a 10-point questionnaire with five possible answers each. The answers range from strongly disagree to strongly agree on a five point scale (strongly disagree being 1 and strongly agree being 5). The items in the questionnaire are (Sauro 2011):

1. I think that I would like to use this system frequently.
2. I found the system unnecessarily complex.
3. I thought the system was easy to use.
4. I think that I would need the support of a technical person to be able to use this system.
5. I found the various functions in this system very well integrated.
6. I thought there was too much inconsistency in this system.
7. I would imagine that most people would learn to use this system very quickly.
8. I found the system very cumbersome to use.
9. I felt very confident using the system.
10. I needed to learn a lot of things before I could get going with the system.

Calculating the SUS score happens in the following way. For odd items one is subtracted from the user response. For even items the user response is subtracted from five. All values are then scaled from zero to four. All values are added up and then multiplied by 2,5 which gives a total value from 0 to 100. (Sauro 2011.)

500 studies have shown the average SUS score to be 68. This can be thought as a reference point to describe if the usability of the product under evaluation is above or below average. (Sauro 2011.)

Net promoter score (NPS)

Net promoter score illustrates customer satisfaction and loyalty in one simple metric. The user is asked only one question to determine the NPS: How likely is it that you will recommend this product to a friend or colleague? The question is answered on a scale of zero to ten (Figure 13). (Meyer 2019.)

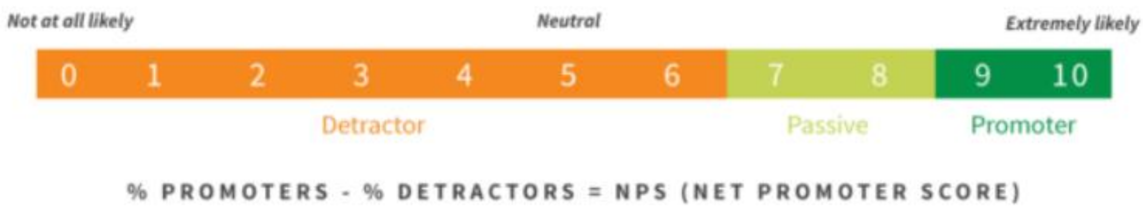


Figure 13: Net promoter score (Meyer 2019).

The answers are grouped into three categories (detractors 0-6, passives 7-8, and promoters 9-10). When calculating the score the passives are not taken into account in the calculation. NPS is calculated as % promoters - % detractors = NPS. (Meyer 2019.)

Customer satisfaction (CSAT)

Customer satisfaction score is a UX metric that expresses customer satisfaction in a convenient metric. Users are asked only one question: How satisfied are you with the product? The result of CSAT is a percentage from 0 to 100 where 100 is the maximum customer satisfaction. Usually the scale of CSAT includes five rating options from very unsatisfied to very satisfied (Figure 14). (Meyer 2019.)

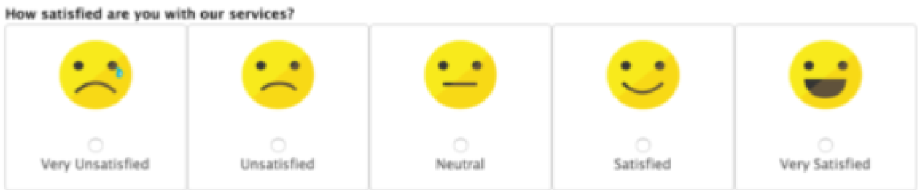


Figure 14: Customer satisfaction scale (Meyer 2019).

The CSAT score is calculated as number of satisfied customers (users who answered either satisfied or very satisfied) divided by total number of respondents and then multiplied with 100% (Meyer 2019).

These seven preliminary UX metrics acted as a baseline for the second round of the research and were put under evaluation there. The total number of UX metrics in the first phase was already quite high since there were also the two additional ones (safety and effectiveness) that are required by the IEC 62366-1 (2015) standard. All the UX metrics and their usefulness had to be reviewed critically during the second phase of the research so that they are not measured only because of the joy of measuring but that they are beneficial for the actual usability engineering and UX design work.

5.3 Second round of research execution – DMAIC process

In the second round of the research and analysis DMAIC process was followed. DMAIC process was selected because it can be used to any business problems and it allows to have an impact on financial results when possible. With the help of DMAIC process the business problem, the content of lean UX canvas, and the preliminary UX metrics were put under closer examination.

A realistic approach to DMAIC process was taken. It meant that only the suitable tools and methods for the business problem were chosen to be used and for example the roles and belt levels (green belt, black belt etc.) were not taken into consideration during the research and analysis. Taking the belt levels into consideration would have been a challenge anyway since there was no project team included in the research performed. The idea was more to utilize the beneficial parts of the process, try it out, and to demonstrate that it can be used also in a business problem like this.

The following chapters go through the five phases of DMAIC process (define, measure, analyze, improve, control). After them the whole process is summed up in a single page A3 report.

5.3.1 Define

In the define phase the problem space was examined closer from the customer's perspective. Important topics that were defined were the environment, problem, critical-to-quality (CTQ) requirements, areas that need development, and objectives. Existing knowledge that has been collected during the past years inside the UX team was utilized in the definitions.

Environment

The environment where the problem takes place is the product development of laboratory automation devices and more precisely the usability engineering and user experience design there. In the product development department a product development process and its sub processes (e.g. usability engineering process) are defined and in use in every project.

Also various stakeholders (internal and external) are included in the processes and their input is collected in different phases of the product development (e.g. information from product management, other research and development teams, partners, and end users).

Problem

Currently the outcome and the success of usability engineering and user experience design are not measured as efficiently as they could be. Without proper measuring it is basically impossible to prove that the UX activities performed have been worthwhile and effective. The devices are released to market when they pass the summative usability evaluation. In the summative usability evaluation it is ensured that the use of the devices is safe and effective, but safety and effectiveness do not reveal e.g. user's level of satisfaction.

The resources available for usability and user experience work are limited in the busy project based working style. It is not possible or even sensible to invest resources in everything. How could the success be properly measured to help realize where to allocate the limited UX design resources that are available?

Although it should not be the case, occasionally the importance of UX design still needs to be justified. Being able to clearly show the benefits of UX design can help changing the mindsets of people in a more positive direction.

CTQ requirements

Figure 15 presents the critical-to-quality requirements in a tree diagram. It consists of three levels that get more specific on each level. First level is the customer needs, second level the factors that affect the customer needs, and the third level the actual CTQ requirements.

The CTQ requirements tree diagram was built based on user research findings that have been collected during the past few years at Thermo Fisher Scientific. The starting point for the CTQ requirements was the customer need for a functioning laboratory automation device that fulfills its intended purpose. This need is affected by use safety, high uptime, and ability to use the device without problems.

From the factors of use safety, high uptime, and ability to use the device without problems the CTQ requirements can be derived. Functioning safety features and lack of hazardous use situations originate from use safety, lack of errors and reliability from high uptime, and usability and user experience from ability to use the device without problems.

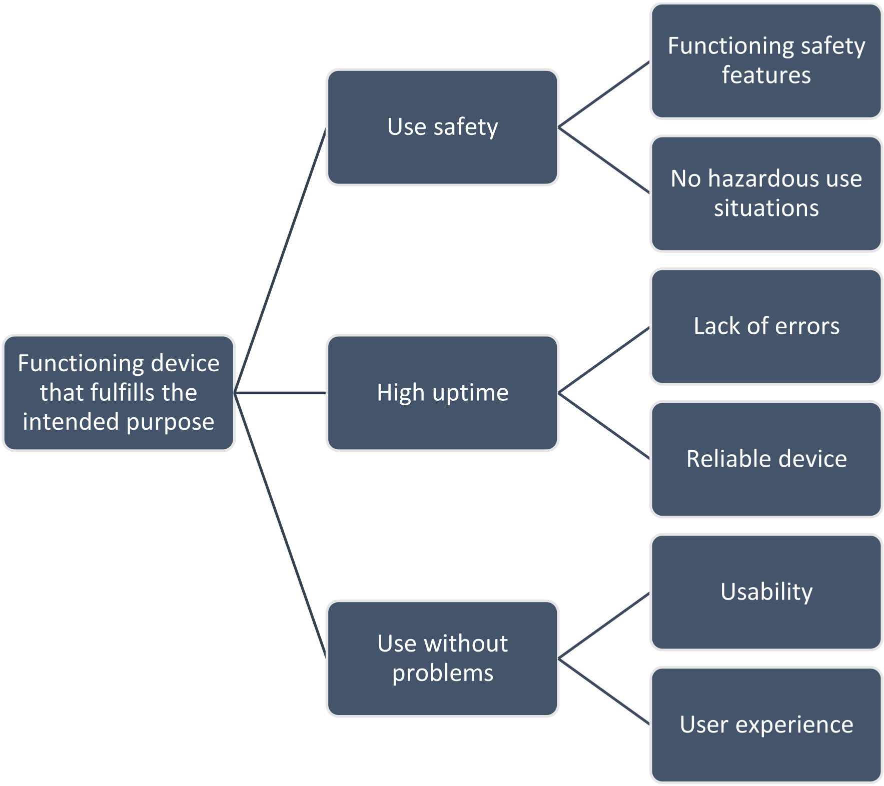


Figure 15: CTQ tree diagram.

Areas that need development

There are basically two areas that need development. First one is the measuring of outcome and success of usability engineering and user experience design. The second one is about making clear in which areas especially to focus on in the usability engineering and user experience design and where to allocate the resources.

Objectives

The main objective is to create a suitable set of UX metrics that work in the design and development of laboratory automation devices. With the help of the UX metrics it is easier to decide where to focus in the usability engineering and user experience design in the future projects. Also justifying the importance of usability engineering and user experience design can be done when the outcome and success can be effectively measured and comparison to earlier measurements can be made.

5.3.2 Measure

In the measure phase SIPOC (suppliers, inputs, process, outputs, customers) diagram and value stream mapping were used to better understand the current processes and to find out potential causes for the problem.

The SIPOC diagram (Figure 16) maps the high level process starting from the UX team and going all the way to the end users of the laboratory automation devices. Customer requirements act as an input for the UX design process and the output of the process is a working laboratory automation device that will be used by laboratory analysts and laboratory technicians. The UX design process is presented in a more detail in the value stream map.

Figure 16: SIPOC diagram for UX design.

Figure 17 shows a simplified value stream map from the usability engineering and UX design point of view. In the value stream map LT stands for lead time and PT for process time. From the lead time and process time it can be seen that UX team's effort is not required for the whole time which allows the team members to work on multiple projects at the same time.

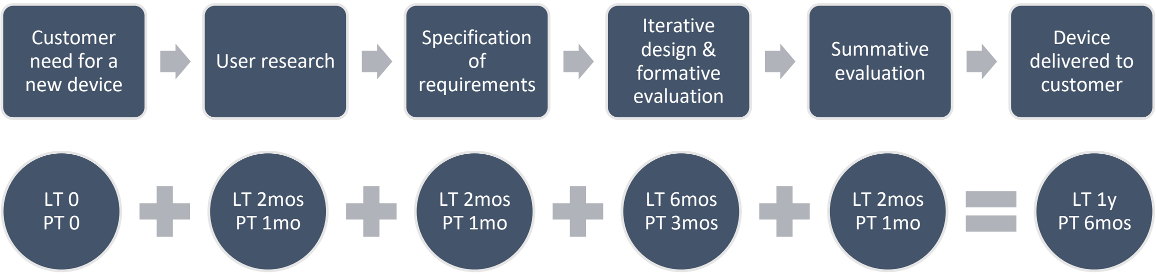


Figure 17: Value stream map from usability engineering and UX design point of view.

As the value stream map shows the product development of a new device starts from a customer need. The process then continues with user research and specification of requirements. Most of the time goes to the iterative design activities and to the formative usability evaluations that are performed throughout the design and development phases. When the project is coming to an end summative usability evaluation is performed before the device can be released to the customer.

The measurements currently in use in the process are safety and effectiveness. Measuring them is mandatory since they both come from IEC 62366-1 (2015) standard that is followed in the usability engineering and user experience design. Safety comes from the fact that the devices must fulfill certain safety standards. Effectiveness on the other hand means that the users must be able to perform certain tasks that belong to the intended purpose of the device. Both safety and effectiveness reflect the quality experienced by the customers.

When the process presented in the SIPOC diagram and in the value stream map, and the current measurements in use were analyzed, two potential causes for the problem were recognized. The current measuring of the results is not as effective as it could since only safety and effectiveness are measured. In addition to safety and effectiveness also new measurements should be taken into use. A suitable set of UX metrics will help in measuring the outcome and success of UX design.

Second potential cause is related to resources. UX team's time is usually not needed for the whole time during the projects which means that the UX team members are working on multiple projects at the same time. Projects have their own schedules and there are situations where busy times in the projects arise at the very same moment. Especially during these busy times it is recognized that the UX resources in use are limited.

5.3.3 Analyze

Two potential causes for the problem were recognized in the measure phase:

1. The success of usability engineering and UX design is only measured with use safety and use effectiveness.
2. The resources in use are limited.

Root cause analysis for these two potential causes was performed using the five whys method. The question why was asked as many times as needed to reach the root causes. In the case of the first potential cause why was asked five times and in the case of the second potential cause why was asked six times in order to reach the root causes.

Root cause analysis for the first potential cause

1. The success of usability engineering and UX design is only measured with use safety and use effectiveness.

Why? It is not mandatory to measure other factors than safety and effectiveness according to IEC 62366-1 (2015).

Why? To measure usability and UX in other ways requires planning.

Why? It is unclear what to measure and what is a good enough level.

Why? It is not easy to measure usability and UX.

Why? Usability and UX consist of many features.

Root cause analysis for the second potential cause

2. The resources in use are limited.

Why? It feels that there are not enough personnel and time.

Why? The resources in use are not fully utilized.

Why? It is not completely clear where to concentrate in the user experience design.

Why? There are no metrics that measure the success of the user experience design.

Why? It is not known what the metrics should be.

Why? Customers have not been asked from.

5.3.4 Improve

In the improve phase several Design of Experiments (DOEs) were planned, performed, and analyzed to affect the root causes of the problem in order to reach the objectives. Table 1 shows the proposed solutions for the DOEs and evaluates them through value analysis. The solutions were analyzed based on how much value do they add to the customers and how easy it is to implement them. The idea is that at least the low hanging fruits (the ones that add value and are easy to implement) should be taken care of first.

Table 1: Proposed solutions organized according to their value.

	Little added value to customers	***A lot of added value to customers***
Easy to implement	• Study what UX metrics are used in general. • Ask internal product management what qualities they value.	• Ask the end users what qualities they value.
Time consuming to implement	• Organize focus groups to find out from end users what qualities are valued.	• Organize usability tests to test UX metrics. • Ask the paying customers what qualities they value. • Create a set of suitable UX metrics for laboratory automation. • Test UX metrics in a full length new product development project

Based on the value analysis most of the proposed solutions were designed to be experimented. The implementation plan for the chosen experiments is presented in Table 2. In the plan the experiments are presented in the order in which they were implemented. Also the predicted results and the actual results of the experiments are included in the plan. After the plan all the experiments are presented one by one.

From the implementation plan in Table 2 it can be seen that proposed solutions to organize focus groups in order to find out what qualities the end users value, and to test the UX metrics in a full length new product development project were left out. End user focus groups were not organized since end users were decided to be interviewed in one-on-one interviews. In focus groups the opinions of others are often influenced by the most dominating persons. Here the opinions of all of the end users that were interviewed were wanted to be expressed without being influenced by the opinions of others. The reason to leave out the testing of UX metrics in a full length new product development project was time related. Although the results might be of great worth a new product development project lasts for such a long time that in the time frame of this thesis there was no possibility to do it.

Table 2: Implementation plan for the DOEs containing also predicted results and the actual results from the DOEs.

DOE	NOTES	PREDICTED RESULTS	RESULTS
STUDY WHAT UX METRICS ARE USED IN GENERAL.	The preliminary list of UX metrics is presented in chapter 5.1.2.	Suitable metrics can be found to work with laboratory automation devices.	Behavioral: Task success rate, time on task, search vs. navigation, user error rate. Attitudinal: SUS, NPS, CSAT.
ORGANIZE USABILITY TESTS TO TEST UX METRICS.	Organized at the same time with the end user interviews.	Task success rate, time on task, search vs. navigation, user error rate are all suitable to be used with laboratory automation devices.	Task success rate and user error rate turned out to be the most useful behavioral UX metrics.

ASK THE END USERS WHAT QUALITIES THEY VALUE.	Interviews held at the same time with the usability tests.	Qualities: ease of use, colleague recommendations, satisfaction.	Qualities: ease of use, safety, scarcity of errors, reliability.
ASK INTERNAL PRODUCT MANAGEMENT WHAT QUALITIES THEY VALUE.	-	Qualities: ease of use, safety, scarcity of errors, reliability.	Qualities: uptime, modularity, reliability, modernity, visual appearance.
ASK THE PAYING CUSTOMERS WHAT QUALITIES THEY VALUE.	Product management interviews revealed lists of tender requirements from customers. They were used to replace the interviews of the paying customers.	-	Qualities: uptime, ease of use of the user features.
CREATE A SET OF SUITABLE UX METRICS FOR LABORATORY AUTOMATION PRODUCT DEVELOPMENT.	-	Behavioral: Task success rate, time on task, search vs. navigation, user error rate. Attitudinal: SUS, NPS, CSAT.	Behavioral: Task success rate (effectiveness), user error rate, safety. Attitudinal: SUS, CSAT, aesthetics, UUA, safety.

DOE #1 – Study what UX metrics are used in general

Already on the first round of the research a preliminary list of seven UX metrics was collected. More information of the preliminary UX metrics and their selection can be seen in chapter 5.1.2 of this thesis.

Those seven UX metrics were chosen because they work for any type of projects (containing both software and hardware), they can be connected to many common business goals, they have an impact on many design decisions, they include both behavioral and attitudinal metrics, they cover the factors that were found important in the discussions inside the UX team, and they were also cited in other sources. The preliminary list of UX metrics contained four behavioral UX metrics and three attitudinal UX metrics. The behavioral UX metrics were task success rate, time on task, search vs. navigation, and user error rate. The attitudinal UX metrics selected were system usability scale (SUS), net promoter score (NPS), and customer satisfaction (CSAT).

These seven UX metrics formed a baseline and they were used as a basis when the following DOEs were planned. In the following DOEs the suitability of the seven UX metrics in laboratory automation use was assessed.

DOE #2 – Organize usability tests

Planning the experiment:

In order to test out the suitability of the behavioral UX metrics (task success rate, time on task, search vs. navigation, and user error rate) usability tests were planned and organized. The behavioral UX metrics were planned to be observed as a part of a summative usability evaluation of a laboratory automation input device.

Altogether five test participants were recruited in the usability testing. The amount of test participants was decided to be five since according to Nielsen (2012) with five test participants almost always the maximum benefit-cost ratio of usability testing is reached. There is little additional benefit when the test is run with more than five people which means that the ROI is dropping fast if additional users are recruited to the usability testing.

The five test participants all represented the intended end users of the laboratory automation devices. All the test participants either worked or had worked in laboratories and they had between one and twenty years of experience from laboratory environment. The test sessions were planned to be held in three separate days and one test session per participant lasted 1-2 hours. More information about the background of the test participants and the shortened versions of the test case descriptions can be found from Appendix 2.

Conducting the experiment and analysis of the experiment results:

The test participants performed individually the test cases presented in Appendix 2 and at the same time their performance was observed. Extra attention was put in observing the behavioral UX metrics from the preliminary list of UX metrics. During the test sessions the suitability of the behavioral UX metrics was reviewed together with safety and effectiveness.

As a result some of the UX metrics seemed to work better than others. Here below the suitability of the behavioral UX metrics from the preliminary list of UX metrics, safety, and effectiveness are analyzed one by one based on the observations made in the usability testing.

Task success rate

Task success rate proved to be a useful metric based on the usability testing. All the test cases had a clearly defined endpoint so it was easy to measure the task success rate. Now when the usability testing was performed as a summative usability evaluation it was required that all the test participants were able to perform all the test tasks. For medical devices no exceptions are allowed and in the testing the task success rate was 100%.

Task success rate turned out to be actually the same metric as effectiveness that is already measured every time in the summative usability evaluation. It is anyway a useful metric that could be taken into wider use also during formative usability testing.

Time on task

During the usability testing the test cases are prepared beforehand and the one hour long test session might actually represent e.g. for one month of usage. For this reason measuring the time on task might not be that useful and is not recommended to be measured. As an exception it could anyway be used to measure for how long it takes to solve a certain problem situation. In the laboratories the time is though greatly affected by the time that it takes to notice the problem. Often this time is longer than the time that is actually needed to solve the problem.

Search vs. navigation

The graphical user interface of the input device did not contain search function so this metric could not be measured. Currently this is the case also with many other laboratory automation devices so having search vs. navigation as metric is not useful. If in the future the graphical user interfaces become more complex and search function is added this metric can be reconsidered.

User error rate

User error rate proved to work well in the test situation. Although it does not reveal the reasons for the errors it clearly shows which test cases are the most challenging ones for the users. To these challenging areas it could be a good idea to invest more design resources.

During the usability testing it could be seen (Table 3) that test case 1 (see Appendix 2 for the test case descriptions) was clearly the most challenging one for the users although at the end all users were able to perform all of the test cases. Test case 1 was related to starting the system which is a task that contains many phases. Starting the system is also something that is done quite rarely since normally the laboratory automation system is running for most of the time. In this case the results suggest that as a follow up it could be a good idea to test the learnability of this feature.

Table 3: Error occurrence rate during usability testing.

TEST CASE	HOW MANY USERS MADE ERROS?	ERROR OCCURRENCE RATE
1	4 / 5	80 %
2	1 / 5	20 %
3	1 / 5	20 %
4	2 / 5	40 %
5	0 / 5	0 %
6	2 / 5	40 %

Safety

During the test sessions no hazardous situations were observed. In the summative usability evaluation the device under evaluation must be already safe to use. Otherwise it cannot make its way all the way to the summative usability evaluation. It is anyway important that the users also feel safe when using the device. For this reason it is recommended to ask the users to rate how safe they felt themselves and the device to be while using it.

All the test participants were asked to give a rating on a scale of zero to ten to rate the use safety. The average rating among the participants was 9,4. This rating and the fact that no hazardous situations were observed suggest that the use safety of the device was on a very good level.

DOE #3 – Ask the end users what qualities they value

Planning the experiment:

At the same time when the usability testing was performed the same test participants were also planned to be interviewed. The same users were used for the interviews due to the convenience that they were already at the company premises and because recruiting even the five users from the actual intended user group turned out to be challenging enough. It was anyway estimated that the amount of interviewees is suitable to give eligible results.

These end user interviews were planned to be conducted as semi-structured one-on-one interviews. End user focus groups were not wanted to be organized since quite often in focus groups the opinions of others are influenced by the most dominating persons. Here the opinions of all of the end users were wanted to be expressed without being influenced by the opinions of others.

The questions for the interviews were prepared beforehand. The structure of the interviews was planned to follow the questions, but in a way that the conversation can be followed and clarifying questions asked when needed. Under specific examination were the preliminary listed attitudinal UX metrics and the qualities behind them (ease of use, colleague referrals and recommendations, and customer satisfaction).

Conducting the experiment and analysis of the experiment results:

After the usability test sessions four out of the five test participants were interviewed. One test participant had to leave early and was not able to attend the interview. The interview questions and the notes from the four interviews can be seen in Appendix 3.

In the interviews it was examined how important the usability, colleague referrals, and general satisfaction actually are. The interviews revealed also other qualities that were considered especially important.

All of the interviewees valued ease of use highly which suggests the importance of system usability scale as a metric. Other qualities that were appreciated contained e.g. safety, scarcity of errors, and reliability. The interviewees also thought that the devices should be simple to use.

Colleague referrals on the other hand did not play a big role. Usually the people at the laboratories are in contact with the people from the same laboratory and knowledge about devices of other manufacturers come mainly from their previous work places.

The interviews revealed that general satisfaction is a quite wide term and the use of it as a metric is a bit questionable. It gives a good general understanding of the level of satisfaction but does not tell why the users are either satisfied or dissatisfied. Asking the level of satisfaction is anyway fairly easy task to do.

DOE #4 – Ask internal product management what qualities they value

Planning the experiment:

Since the product management department of laboratory automation devices is rather small there were only three interviews (all relevant positions from product management) planned. The people interviewed had altogether dozens of years of experience from the laboratory automation, knowledge about the field in general, and knowledge about customers and their needs and requirements.

The three product management interviews were planned as semi-structured one-on-one interviews. One-on-one interviews were chosen in order to see if there would be any differences between the opinions of interviewees. The interviews were planned to be conducted after the end user interviews so that the findings would not affect the question setting of the end user interviews.

There were quite a few questions prepared beforehand about the topics related to tender requirements, customer wishes, improvement ideas for current devices, and product qualities. The idea was to keep the conversation flowing and not to stress if all of the topics are not covered in all of the interviews. After all the people working in the product management had different emphases in their job roles.

Conducting the experiment and analysis of the experiment results:

All of the interviews performed did not include all of the questions prepared. The discussions were flowing quite freely based on the topics but everything needed got covered. The interview questions and the notes can be found from Appendix 4.

In the interviews the main idea was to find out what qualities are valued by the product management and based on their understanding what are the qualities valued by the customers. The valued qualities were mainly related to product features, costs, and modularity of the system but there were also qualities that can be affected by usability engineering and user experience design.

One of the most important qualities was uptime and everything related to that. The laboratory automation system should be running close to 100 % of the time or otherwise the laboratory can have massive issues. Long stops are also bad for company reputation and might have a negative effect to the future sales.

The qualities valued by product management differed quite a lot from the qualities valued by the end users. End users valued ease of use and safety, whereas product management was more interested in reliability, modernity, and visual appearance of the devices. They all were qualities that can be noticed from the devices quite easily and can possibly increase the sales of the devices.

DOE #5 – Ask the paying customers what qualities they value

Planning the experiment:

The original idea for the experiment was to identify paying customers and interview them in order to find out what qualities they value. This turned out to be a challenging task since the selling process of the laboratory automation devices happens mainly through distributors and there are no direct contacts to paying customers. During the product management interviews it was found out that there is anyway access to customer's tender requirements (see Appendix 5 for the tender requirements) so DOE #5 was redesigned to be:

DOE #5 – Examine what usability and UX qualities can be found from tender requirements.

In the redesigned DOE the plan was to go through the lists of tender requirements, to examine if usability and UX related qualities can be found from them, and to analyze if the qualities can be utilized when thinking about UX metrics.

Conducting the experiment and analysis of the experiment results:

Already during the product management interviews it was found out that the tender requirements were mainly technical or related to certain features that the system must have. This turned out to be true when the tender requirements were examined.

The tender requirements contained mainly technical aspects but they had anyway requirements that can be at least indirectly affected with usability engineering and user experience design activities and measured with certain UX metrics. This kind of requirements included user features that are used from the graphical user interface and uptime.

DOE #6 – Create a set of suitable UX metrics for laboratory automation product development

Planning the experiment:

Based on the results and analyses of the previous experiments and preliminary list of UX metrics a set of suitable UX metrics for laboratory automation product development is created.

Conducting the experiment and analysis of the experiment results:

Based on all previously mentioned preliminary list of UX metrics, usability tests, end user interviews, product management interviews, and tender requirements, it was possible to create a set of UX metrics that is suitable for laboratory automation product development. The list consists of the following UX metrics that can be divided into behavioral and attitudinal UX metrics (a more detailed description of the recommended UX metrics is provided in chapter 6.1):

- Behavioral
  - Task success rate / Effectiveness
  - User error rate
  - Safety
- Attitudinal
  - System usability scale
  - Customer satisfaction
  - Aesthetics
  - Utility, usability, and aesthetics
  - Safety

Out of the behavioral UX metrics effectiveness and safety are already required by IEC 62366-1 (2015) standard but measuring them got support from the results of usability testing and end user interviews too. The importance of task success rate and user error rate were also justified with the results of usability testing and end user interviews. Task success rate and user error rate are very significant metrics what comes to uptime, which was seen as one of the most important qualities in the product management interviews and tender requirements. All the behavioral UX metrics can be measured by usability testing.

From the attitudinal UX metrics system usability scale and safety were seen as important metrics since ease of use and use safety were top product qualities according to the end users. Customer satisfaction and aesthetics on the other hand were something that were valued by the product management and based on their understanding by the paying customers too. Measuring users' perception about satisfaction and aesthetics is also quite straightforward. Utility, usability, and aesthetics is a UX metric that can be used to replace system usability score, customer satisfaction, and aesthetics as individual metrics, or it can be used in addition to them. Utility, usability, and aesthetics survey contains only three statements so it can be used to calculate a total UX rating also in busy situations where there is not a lot of time available. All the attitudinal UX metrics can be measured by asking from the users or by creating a survey.

5.3.5 Control

UX metrics suitable for laboratory automation device product development were successfully created based on metrics used in other fields of industries, usability testing, and interviews with various stakeholders. Based on the end user and product management interviews and after trying out some of the metrics in the usability testing, it can be assumed that the selected metrics work as intended.

The measures taken have succeeded in creating metrics for usability engineering and user experience design and thus the objectives initially created can be considered as achieved.

A separate control plan was not written in the context of this thesis. In order for the improvements not to fade away there are actions that should be done as a follow up. First of all the recommended UX metrics need to be piloted in an upcoming new product development project. The usability engineering process should be updated to contain the new UX metrics that support the design work. When the improved process is transferred back to the process owner the documented process and its improvements and benefits must be communicated to the process owner. After the usability engineering process is updated also appropriate document templates should be updated to include the new UX metrics. Also communicating the findings and improvement to the whole UX team is an important action that must not be forgotten.

5.3.6 A3 report

The whole DMAIC process is summed up in a one single page A3 report. The A3 report can be seen in Appendix 6.

6 RESEARCH FINDINGS AND RECOMMENDATIONS

The sixth chapter of the thesis explains the findings of the research that was performed. It presents the recommended UX metrics to be used to measure the outcome and success of usability engineering and user experience design, and gives answers to the research questions that were formed in the beginning of the thesis.

6.1 Recommended UX metrics

Through theory and practice it was clear that UX metrics have a significant role when measuring the outcome and success of the usability engineering and user experience design. If the user experience is not measured, how can it be known how the UX is performing and if design changes made are for the better or for the worse (Turner 2018)? The UX metrics are also commonly used to communicate the business impact of the UX design and to convince management about the benefits of usability engineering and user experience design.

Based on preliminary list of UX metrics, usability tests, end user interviews, product management interviews, and tender requirements a list of UX metrics suitable for laboratory automation product development was created. In the research theory and practice were combined to find metrics that can be efficiently utilized when developing laboratory automation devices.

When looking at the recommended UX metrics they might look quite general and that they could be used in any field of industry. This is partly true since nothing prevents from using any of the UX metrics in any context. The important part was to find UX metrics that are the most useful ones in the specified context. The UX metrics recommended here were put under review in various experiments to prove that they work especially well in laboratory automation product development. The recommended UX metrics can be used as tools that enable performing measurements in the daily work and prove the benefits of usability engineering and UX design.

The recommended UX metrics measure qualities that were found important by the end users and product management, and that have an effect on the requirements required in tenders. They take into account the monetary values in the form of uptime and contain both behavioral and attitudinal UX metrics. They can also be used to guide the resource allocation and to justify the importance of usability engineering and user experience design. The recommended UX metrics are the following:

- Behavioral UX metrics
    - Task success rate / Effectiveness
    - User error rate
    - Safety
- Attitudinal UX metrics
    - System usability scale
    - Customer satisfaction
    - Aesthetics
    - Utility, usability, and aesthetics
    - Safety

Task success rate / Effectiveness

Task success rate turned out to be the same metric as effectiveness which is required to be measured when IEC 62366-1 (2015) standard is followed. For this reason and based on the results of the experiments performed it is natural that the metric of task success rate / effectiveness is in the list of recommended UX metrics.

Task success rate can be used whenever a task has a clearly defined endpoint. It measures the number of correctly executed tasks compared to the total number of attempts. It does not give information why a user could not successfully complete a task but it is anyway a very valuable indicator. (Meyer 2019.) Task success rate can be used during formative and summative usability testing.

The task success rate is presented as a percentage. If for example nine out of ten users are able to complete a task, the task success rate is calculated in the following way:

9 / 10 = 90 %

User error rate

User error rate was proved to work well in the usability testing and it is also an indicator to examine the uptime of the laboratory automation system. Although it does not reveal the reasons for the errors it clearly shows which tasks are the most challenging ones for the users. To these areas it could be a good idea to invest more design resources.

User error rate describes the number of times a user made an error while working on a task. It gives an idea of how clear and user friendly the product is. The higher the user error rate score is, the higher is the number of usability problems. (Meyer 2019.)

There are two frequently used ways to calculate the user error rate: error occurrence rate and error rate. The error occurrence rate is used if a task only allows one potential error or if there are several but only one of them is wanted to be measured. It is calculated as the total number of errors that occurred for all users divided by the total number of possible errors for all users. Error rate on the other hand can be used if multiple errors per tasks are possible or if multiple errors are wanted to be measured. In error rate the number of errors is divided by the total number of attempts. (Meyer 2019.)

Safety

Safety was seen as one of the most important qualities among the end users. For them the use of the devices in the laboratory must be safe. In addition to effectiveness also safety comes from the IEC 62366-1 (2015) standard so measuring it is a mandatory thing to do.

Safety is categorized as both behavioral and attitudinal in the recommended UX metrics. Whenever usability testing is performed it should be taken care of beforehand that using the device is safe for the user, for the device, and for the environment. It can anyway be observed if during usability testing any of the user actions might lead to a hazardous situation and for this reason safety is categorized as a behavioral UX metric.

It is also important that users feel safe when using the device. After testing the device users can be asked to rate how safe they felt themselves and the device to be while using it. For this reason safety can also be categorized as an attitudinal UX metric.

System usability scale (SUS)

Ease of use was ranked as the number one quality by the end users. They also appreciated scarcity of errors and thought that the devices should be simple to use. These all suggest the importance of usability of the devices.

System usability scale is quick tool for testing the usability of a product. It consists of a 10-point questionnaire with five possible answers each. The answers range from strongly disagree to strongly agree. Based on the answers a total value in the range of 0-100 can be calculated for the product. (Sauro 2011.) The higher the rating, the better the usability.

500 studies have shown the average SUS score to be 68. This can be thought as a reference point to describe if the usability of the product under evaluation is above or below average. (Sauro 2011.)

Customer satisfaction (CSAT)

Using customer satisfaction as a metric gives a good understanding of the general level of satisfaction but does not tell why the users are either satisfied or dissatisfied. Asking the users to rate the level of satisfaction is anyway an easy task to do and does not require a lot of resources.

Users are asked only one question: How satisfied are you with the product? The result of CSAT is a percentage from 0 to 100 where 100 is the maximum customer satisfaction. Usually the scale of CSAT includes five rating options ranging from very unsatisfied to very satisfied. (Meyer 2019.)

Aesthetics

Aesthetics and visual appearance emerged as important qualities in the product management interviews since they are the first things that are noticed from the devices and might work as qualities that attract customers.

Aesthetics can be asked simply by using one statement and a 0 to 5 scale from strongly disagree to strongly agree. The statement to be used is: This device is aesthetically appealing and appropriate. (Myles 2019.)

The statement focuses on the look and feel of the product. It asks the users to judge if the device is appealing to them and if its style is appropriate for the context. (Myles 2019.)

Utility, usability, and aesthetics (UUA)

Utility, usability, and aesthetics survey can be used in addition to or instead of the attitudinal UX metrics of system usability scale, customer satisfaction, and aesthetics. UUA survey is a method to calculate a total UX rating and it can be used also in situations where there is no time run large, time consuming questionnaires or surveys.

UUA survey consists of only three statements. Utility is measured with the statement “this device’s capabilities meet my requirements”, usability with the statement “this device is easy to use”, and aesthetics with the statement “this device is aesthetically appealing and appropriate”. (Myles 2019.)

All statements are scored on 0 to 5 scale, where 0 means strongly disagree and 5 strongly agree. This gives a clear indication of user perception of utility, usability, and aesthetics. UUA survey gives different weighting for the statements. Utility is worth 3x, usability 2x, and aesthetics 1x (Figure 18). (Myles 2019.) This gives the following equation for the total UX rating:

UX rating = ((Utility x 3) + (Usability x 2) + Aesthetics) / 6

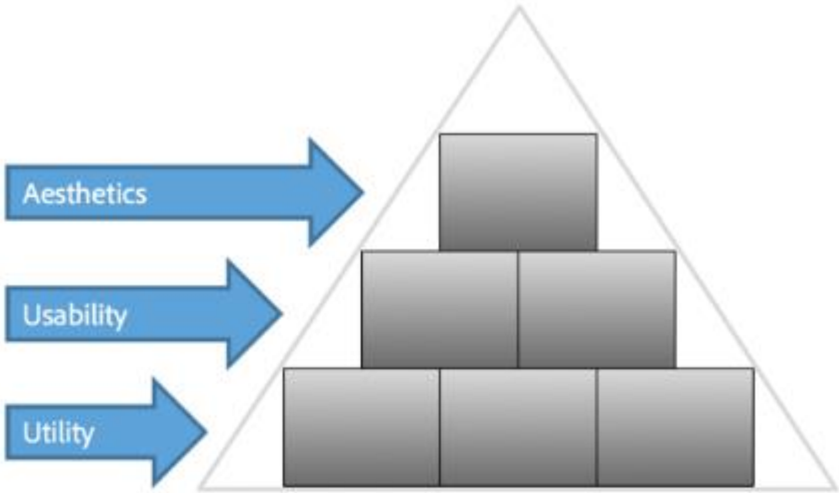


Figure 18: Utility, usability, and aesthetics pyramid (Myles 2019).

Figure 19 shows target ranges for the total UX rating. The ranges are categorized as follows (Myles 2019):

- Poor (<= 2): Ratings in this range need further investigation to resolve the root cause for such a low score.
- Needs work (2,1 – 3,4): Not negative but not positive either.
- Good (3,5 – 4,4): Scores above 3,5 are results that are wanted to be seen.
- Excellent (4,5 – 5): Exceptional user satisfaction.

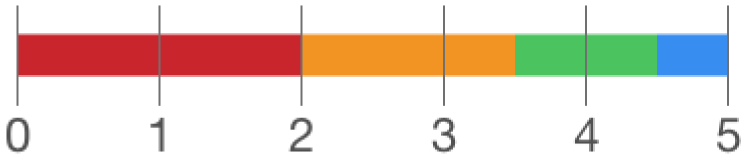


Figure 19: Target ranges for UUA scores (Myles 2019).

6.2 Answers to research questions

The four research questions of the thesis were defined in the first chapter. After getting acquainted with the research methodologies through theory, collecting the data, performing the experiments, and analyzing the results it is time to answer the research questions. When answering the research questions it is good to remember that the answers are based on empirical research performed as well as on the theory presented in this thesis.

UX metrics: What UX metrics can be efficiently utilized when developing laboratory automation devices?

Through theory and practice it became clear that UX metrics have a significant role when measuring the outcome and success of the usability engineering and user experience design. If the UX is not measured, how can it be known how the UX is performing and how to know if design changes made are for the better or for the worse (Turner 2018)? The UX metrics are also commonly used to communicate the business impact of the UX design and to convince management about the benefits of usability engineering and user experience design.

A suitable set of UX metrics for laboratory automation product development was successfully created. In the research theory and practice were combined to find metrics that can be efficiently utilized when developing laboratory automation devices. The UX metrics recommended here were put under review in various experiments to prove that they work in their intended context. The final, recommended list of UX metrics takes into account the monetary values in the form of uptime and contains both behavioral and attitudinal UX metrics. The UX metrics can also be used to guide the resource allocation and to justify the importance of usability engineering and user experience design. The recommended list contains the following UX metrics:

- Behavioral UX metrics
    - Task success rate / Effectiveness
    - User error rate
    - Safety
- Attitudinal UX metrics
    - System usability scale
    - Customer satisfaction
    - Aesthetics
    - Utility, usability, and aesthetics
    - Safety

Return on investment: How can the ROI of UX activities be measured?

Measuring directly the ROI of certain UX activities is challenging because the ROI might not be income that is generated. The ROI of UX activities is more about reducing the costs by eliminating poor design. During the experiments performed in the research phase it was made clear that the downtime of laboratory automation systems is causing a lot of costs and it should be avoided by any means. Usability engineering and user experience activities that affect directly in keeping the uptime high can be seen to bring monetary value. Qualities that help keeping the uptime high are e.g. ease of use and lack of errors.

Usability testing is one preferred way to increase the ROI of UX activities since due to usability testing reduced training, maintenance, production, and revision costs can be seen. When organizing usability tests the goal should be in reaching a maximum benefit-cost ratio. According to Nielsen (2012) recruiting five test participants is enough since there is only a little additional benefit when the test is run with more than five people. At the same time recruiting additional users to the usability tests means that the ROI is dropping fast.

The ROI of UX activities can be categorized into soft and hard dollars based on the benefits. The soft side includes results such as increased customer loyalty, increased word-of-mouth referrals, increased efficiency, decreased user errors, and productivity gains. On the hard side there are increased sales, cost savings from support and training, less money spent on development, and less money spent on redo cases. (Philips 2018.)

Stakeholder information: How to align the internal and external stakeholder information?

In the design and development of laboratory automation devices multiple stakeholders are involved. Both the internal and external stakeholders have their own needs and requirements for the end products. Aligning the internal and external stakeholder information becomes important especially when design resources are limited.

During this thesis the opinions of product management (internal stakeholders) and end users (external stakeholders) were listened to. It was learned that actually there was no conflict between these opinions. They shared common interests but in general the product management was more interested in qualities that can directly boost the sales of the laboratory automation devices, whereas the end users were more interested in qualities that affect the use of the devices. This is only natural since the end users are the ones who are using the devices on a daily basis.

Future: How can the UX metrics be utilized in coming projects?

In order to help bringing extra value, the UX metrics must be taken into account in the future projects. For the improvements not to fade away there are actions that need to be done as a follow up. First of all the recommended UX metrics need to be piloted in an upcoming new product development project. The usability engineering process should be updated to contain the new UX metrics that support the design work. When the improved process is transferred back to the process owner the documented process and its improvements and benefits must be communicated to the process owner. After the usability engineering process is updated also appropriate document templates should be updated to include the new UX metrics. Also communicating the findings and improvement to the whole UX team is an important action that must not be forgotten.

6.3 Reliability and validity

The research performed in the thesis was based on a real life problem that was observed in the daily work. Since the recommendations are wanted to be taken into use also their reliability and validity should be evaluated in order to figure out the quality of the research.

The used data collection methods included e.g. usability testing and interviews. Out of the data collection methods interviews were the most used ones. The interviews were all hold as one-on-one interviews so that the opinions of the interviewees would not be affected by others and biased. All the data collected during the research phase has been stored and analyzed on the way. All the findings and recommendations were made based on the data and analysis, and they can be traced back to the data. Being able to follow the path backwards suggests that the same results would be got if the research was to be repeated. This is the basis of reliability and suggests that the findings and recommendations made in the thesis are reliable.

It can be speculated whether additional interviewees would have added value for the results of the thesis or whether the findings and recommendations would have been the same in that case. What comes to the product management interviews, the company opinions were reached since all the relevant persons from product management were interviewed. In the end user interviews the same topics and opinions started to be repeated which was the case also in the usability testing. In usability testing it has been studied that testing with more than five users does not result in appreciably more insights (Nielsen 2012). The data collected from the different sources (end user interviews, product management interviews, usability testing, tender requirements) revealed similar findings which suggest that right topics were researched and that the research performed was valid. Also the final recommendations were made only after all the data was collected and findings analyzed to avoid assumptions that would be made too early.

7 CONCLUSIONS

This thesis studied measuring the outcome and success of usability engineering and user experience design in laboratory automation product development by utilizing UX metrics. As an outcome of the thesis a set of UX metrics suitable for laboratory automation product development is recommended to be used. When looking at the recommended UX metrics they might look quite general and it may feel that they could be used in any field of industry. This is partly true since nothing prevents from using any of the UX metrics in any context wanted. The important part was to find UX metrics that are the most useful ones in the specified context. The UX metrics recommended here were put under review in various experiments to prove that they work especially well in laboratory automation product development. They can be used as tools to enable measuring in the daily work.

The findings presented in this thesis are based on the theory explained as well as on the experiments and observations made during the empirical research phase. The recommended UX metrics work in measuring the success of usability engineering and user experience design in laboratory automation product development and they can be used to guide the resource allocation when needed. In addition to that the measurements that are made using the UX metrics can be used to justify the importance of usability engineering and user experience design if needed.

From early on it was clear that the topic of the thesis is related to the UX metrics and measuring the success of usability engineering and user experience design. The setting of the objectives of the thesis and limiting the topic became more precise as the work progressed. In order to give precise limits for the thesis it was decided to limit the topic to cover only the product development of laboratory automation devices manufactured by Thermo Fisher Scientific. Other product lines and the devices of other manufacturers were excluded from the scope of the thesis. This helped in designing the experiments that were performed during the research phase.

The theoretical part of the thesis included relevant information related to the theoretical background of the work and to the research methodologies that were used. The discussion of the theoretical background started with the definition of basic concepts such as user centered design, usability, and user experience and included also the principles of UX metrics and the ROI of UX activities. The aim was to keep the theoretical background rather short, but still relatively comprehensive, so that also employees involved in user interface design who are slightly less familiar with the topic could get acquainted with the content of the thesis. The theoretical part also presented the research methodologies (such as Lean Six Sigma) that were relevant for the empirical part of the thesis.

Lean UX canvas and DMAIC process were selected as the main tools and techniques for the empirical research. The lean UX canvas was selected because it could be used to address the business problem, to find solutions to the problem, and to see towards which direction to move next. DMAIC process on the other hand offered a framework and tools to examine the business problem closer, to analyze the root causes of the business problem, and to find solutions for it. Using the lean UX canvas gave a good starting point for the research. It helped in structuring the first thoughts into one canvas where everything could be easily addressed. Following the phases of DMAIC process made sure that all necessary aspects within the business problem got examined. Using DMAIC process worked well in the context since only the tools and techniques that were beneficial for reaching the objectives were decided to be used.

The most important outcome of the thesis was the list of recommended UX metrics. During the research work, it was found out that there is a large variety of UX metrics available but many existing UX metrics as such are not suitable as indicators for the product development of laboratory automation devices. Therefore, suitable UX metrics for usability engineering and user experience design of laboratory automation product development had to be sought. Not all metrics presented and recommended in the thesis were evaluated in practice in the usability tests. However, this is not considered very meaningful for the results of the work, as the metrics were comprehensively studied through theory and partly also through practice. Arguments for all the recommended UX metrics are seen in the results of the experiments.

At the beginning of the thesis, four research questions were identified. During the research and analysis answers to these questions were sought. The questions were related to the main topics of the work, i.e. UX metrics, ROI of UX activities, stakeholder information, and future utilization of the UX metrics. These questions guided the progress of the research and prevented the work from expanding excessively. During the work, the answers to all four research questions were found.

During the work it was noticed that usability and user experience are important especially for the end users but the end users are not the ones making the buying decisions of the laboratory automation devices. Through usability engineering and user experience design it is easier to have an effect on the needs and wishes of the end users than on the requirements of the paying customers. Usability and user experience qualities are not directly part of tender requirements but they give value for their part especially by having an effect on the uptime of the system. High uptime is one of the most important requirements for the laboratory automation since in some cases laboratories are expected to be running

24/7. Every minute the system is not running means extra costs. This leads to the conclusion that the ROI of UX activities is more about reducing the costs by eliminating poor design than about creating direct income.

Table 4 shows a SWOT analysis of the thesis. The SWOT analysis takes into account the strengths, weaknesses, opportunities, and threats of the thesis. When considering future usability engineering and user experience design, special attention should be paid to the realization of the presented opportunities. At the same time, the likelihood of threats presented can be reduced.

Table 4: SWOT analysis of the thesis.

	+	-
Internal origin	**STRENGTHS** The recommended UX metrics are proven to work especially with laboratory automation product development. ROI of UX activities is proven to be positive.	**WEAKNESSES** Partly rather old references used. Paying customers were not interviewed in the research phase.
External origin	**OPPORTUNITIES** UX metrics can be used to make work practices more effective. UX metrics can help in allocating UX design resources.	**THREATS** Recommended UX metrics are not taken into use. Some of the recommended UX metrics prove to be useless in practice.

There is of course a threat that the recommended UX metrics are not taken into use. Here the opportunities that the UX metrics are offering anyway overcome the threats. When the UX metrics are used to measure the outcome and success of the work there is an opportunity to make the work practices more effective. Since all of the UX metrics were not tested in practice it might be possible that some of the recommended metrics do not work as planned or are seen of no use. If this is the case the preferred option is to drop those metrics and to concentrate on the ones that help bringing more value.

In some places, the references used in this thesis are rather old, so one can think about whether they are still relevant. However, the old reference material involved in this thesis has been selected on purpose. These old references from Nielsen and Norman are the basic books related to usability and their content is considered still even today relevant and valid. Even today these basic books are cited in publications all over the world when usability is discussed. In this thesis there has also been a desire to use the original sources when presenting certain theoretical backgrounds, although the same topics may have been addressed in more recent works or articles. In these cases, it has been considered more important to indicate where the theory originally came from. In return for these rather old references, new articles especially concerning UX metrics and ROI of UX activities are also included.

All in all, the outcome of the thesis can be considered successful. The answers to all of the research questions were found and the research carried out met the objectives set for it. The work succeeded in finding a set of suitable UX metrics for laboratory automation product development and they are ready to be utilized in the future projects. The first impression of the findings of the thesis is that they can be considered valid and reliable. Validity is proven with the fact that the UX metrics created can actually be used for measuring the outcome and success of usability engineering and UX design. Reliability on the other hand is proven with the facts that the UX metrics are proved to work in the intended context through the experiments and that all the findings and recommendations can be traced back to the collected data. It is assumed that additional experiments would only strengthen the suitability of the chosen metrics. When the use of UX metrics is piloted in the future projects it can of course be seen how well they really work in practice. Especially interesting will be the use of UUA survey to calculate a total UX rating for the devices and how the metric can be utilized in the long run. For the improvements not to fade away the next steps recommended would be to communicate the findings of the thesis to the whole UX team, to update the usability engineering process and appropriate document templates to contain information about the recommended UX metrics, and to start using the UX metrics in practice.

REFERENCES

Bevan, N. 2008. *UX, Usability and ISO Standards.* [Accessed 10 October 2020]. Available at: http://www.cs.tut.fi/ihte/CHI08_workshop/papers/Bevan_UXEM_CHI08_06April08.pdf

Bhanushali, N. 2019. *Lean vs Six Sigma - What is the difference between them?.* [Accessed 19 October 2020]. Available at: https://www.greycampus.com/blog/quality-management/six-sigma-vs-lean-six-sigma

Bolderston, A. 2012. Conducting a Research Interview. *Journal of Medical Imaging and Radiation Sciences,* Issue 43, 66-76.

Brue, G. 2015. *Six Sigma for Managers.* 2nd ed. New York: McGraw-Hill Education.

Chen, J. 2020. *Return on Investment (ROI).* [Accessed 25 September 2020]. Available at: https://www.investopedia.com/terms/r/returnoninvestment.asp

Derome, J. 2015. *What is user experience?.* [Accessed 25 September 2020]. Available at: https://www.usertesting.com/blog/what-is-user-experience

Duermyer, R. 2020. *What Is Return on Investment or ROI?.* [Accessed 25 September 2020]. Available at: https://www.thebalancesmb.com/roi-return-on-investment-1794432

Faulkner, X. 2000. *Usability Engineering.* Hampshire: Red Globe Press.

Fried, J. & Heinemeier Hansson, D. 2011. *ReWork.* Porvoo: Atena.

Gothelf, J. 2016. *The Lean UX Canvas.* [Accessed 28 August 2020]. Available at: https://jeffgothelf.com/blog/the-lean-ux-canvas/

Gothelf, J. 2019. *Lean UX Canvas V2.* [Accessed 28 August 2020]. Available at: https://jeffgothelf.com/blog/leanuxcanvas-v2/

Gribbons, W. 2013. *The Four Waves of User-Centered Design.* [Accessed 30 September 2020]. Available at: https://uxmag.com/articles/the-four-waves-of-user-centered-design

Göransson, B., Gulliksen, J. & Boivie, I. 2003. The Usability Design Process - Integrating User-centered Systems Design in the Software Development Process. *Software Process Improvement and Practice,* 8(2), 111-131.

Hassenzahl, M. 2008. *User Experience (UX): Towards an experimental perspective on product quality.* New York, Association for Computer Machinery.

Hill, T. & Westbrook, R. 1997. SWOT analysis: It's time for a product recall. *Long Range Planning,* 30(1), 46-52.

Hyysalo, S. 2009. *Käyttäjä tuotekehityksessä.* Helsinki: Taideteollinen korkeakoulu.

IEC 62366-1. 2015. *Medical devices - Part 1: Application of usability engineering to medical devices.* Geneva: International Electrotechnical Commission.

IEC TR 62366-2. 2016. *Medical devices - Part 2: Guidance on the application of usability engineering to medical devices.* Geneva: International Electrotechnical Commission.

ISO 9241-11. 2018. *Ergonomics of human-system interaction. Part 11: Usability: Definitions and concepts.* Brussels: International Organization for Standardization.

ISO 9241-210. 2019. *Ergonomics of human-system interaction. Part 210: Human-centred design for interactive systems.* Brussels: International Organization for Standardization.

Juti, B. 2016. *Game Changer.* Helsinki: WSOY.

Kenton, W. 2018. *Lean Six Sigma.* [Accessed 20 October 2020].
Available at: https://www.investopedia.com/terms/l/lean-six-sigma.asp

Kolko, J. 2018. *The Divisiveness of Design Thinking.* Interactions, XXV(3), 29-34.

Kuutti, W. 2003. *Käytettävyys, suunnittelu ja arviointi.* Helsinki: Talentum.

Leech, J., 2011. *A big list of UX KPIs and Metrics.* [Accessed 26 October 2020].
Available at: https://www.cxpartners.co.uk/our-thinking/big_list_of_ux_kpis_and_metrics/

Leijala, A. 2015. *1-9 hengen mikroyritykset Suomen kasvun vetureina.* Helsinki: Books on Demand.

Marsh, S. 2018. *User Research: A Practical Guide to Designing Better Products and Services.* London: Kogan Page.

Meyer, S. 2019. *The 7 most important UX KPIs and how to measure them.* [Accessed 13 October 2020].
Available at: https://www.testingtime.com/en/blog/important-ux-kpis/

Moran, K. 2020. *Three Myths About Calculating the ROI of UX.* [Accessed 23 October 2020].
Available at: https://www.nngroup.com/articles/three-myths-roi-ux/

Morville, P. 2004. *User Experience Design.* [Accessed 6 October 2020].
Available at: http://semanticstudios.com/user_experience_design/

Myles, M. 2019. *UX Metrics: Why not measure aesthetics?.* [Accessed 31 October 2020].
Available at: https://uxdesign.cc/measuring-ux-account-for-aesthetics-600fa66f4cd9

Nguyen, T. 2020. *A metrics-driven approach to evaluate success of UX design.* [Accessed 27 October 2020].
Available at: https://uxdesign.cc/a-metrics-driven-approach-to-evaluate-success-of-ux-design-dd3bea098820

Nielsen, J. 1993. *Usability Engineering.* San Diego: Morgan Kaufmann.

Nielsen, J. 2000. *Designing Web Usability.* Indianapolis: New Riders Publishing.

Nielsen, J. 2012. *How Many Test Users in a Usability Study?.* [Accessed 2 November 2020].
Available at: https://www.nngroup.com/articles/how-many-test-users/

Nielsen, J. 2012. *Usability 101: Introduction to Usability.* [Accessed 25 September 2020].
Available at: https://www.nngroup.com/articles/usability-101-introduction-to-usability/

Norman, D. A. 1998. *The Invisible Computer.* Cambridge: MIT Press.

Norman, D. A. 2013. *The Design of Everyday Things.* Revised and expanded ed. New York: Basic Books.

Norman, D. A. 2016. *Living with Complexity.* Gotham: MIT Press.

Osterwalder, A. & Pigneur, Y. 2010. *Business Model Generation.* New Jersey: John Wiley & Sons.

Pacheco, J. 2019. *History and Nature of User-Centred Design.* [Accessed 18 November 2020].
Available at: https://medium.com/@joshuapacheco/history-and-nature-of-user-centred-design-18e20f86294f

Pavliscak, P. 2014. *Choosing the Right Metrics for User Experience.* [Accessed 25 September 2020].
Available at: https://www.uxmatters.com/mt/archives/2014/06/choosing-the-right-metrics-for-user-experience.php

Philips, M. 2018. *The True ROI of UX: B2B Redesign Case Studies.* [Accessed 23 October 2020].
Available at: https://www.toptal.com/designers/ux/roi-of-ux-redesign-case-studies

Plewes, S. & Fraser, J. 2015. *6 Indicators of an Organizations UX Maturity Level – Lessons In UX Maturity: Part 1.* [Accessed 25 September 2020].
Available at: https://www.macadamian.com/learn/indicators-of-user-experience-maturity/

Rastogi, A. 2018. *A Brief Introduction to Lean, Six Sigma and Lean Six Sigma.* [Accessed 20 October 2020].
Available at: http://www.greycampus.com/blog/quality-management/a-brief-introduction-to-lean-and-six-sigma-and-lean-six-sigma

Ratcliff, C. & Kelkar, K. 2019. *What metrics and KPIs do the experts use to measure UX effectiveness?.* [Accessed 13 October 2020].
Available at: https://www.userzoom.com/blog/what-metrics-do-the-experts-use-to-measure-ux-effectiveness/

Rosser, B. 2009. *Better, stronger, faster.* Oxford: Infinite Ideas.

Roto, V., Law, E., Vermeeren, A. & Hoonhout, J. 2011. *User Experience White Paper: Bringing clarity to the concept of user experience.* Dagstuhl, Dagstuhl Seminar on Demarcating User Experience.

Rubin, J. & Chisnell, D. 2008. *Handbook of Usability Testing: How to Plan, Design and Conduct Effective Tests.* 2nd ed. Indianapolis: Wiley Publishing Inc.

Ruusuvuori, J. & Tiittula, L. 2005. *Haastattelu: tutkimus, tilanteet ja vuorovaikutus.* Tampere: Vastapaino.

Sauro, J. 2011. *Measuring Usability with the System Usability Scale (SUS).* [Accessed 27 October 2020].
Available at: https://measuringu.com/sus/

Shneiderman, B. 2003. *Leonardo's Laptop: Human Needs and the New Computing Technologies.* Cambridge: MIT Press.

Spool, J. M. 2017. *A Proven Method For Showing The Value Of Good UX.* [Accessed 23 October 2020].
Available at: https://articles.uie.com/a-proven-method-for-showing-the-value-of-good-ux/

Spool, J. M. 2018. *Yes, Alan, There Is An ROI For UX Design.* [Accessed 23 October 2020].
Available at: https://articles.uie.com/yes-alan-there-is-an-roi-for-ux-design/

Tej, R. 2019. *ROI calculation for UX Design.* [Accessed 23 October 2020].
Available at: https://tejjj.medium.com/roi-calculation-for-ux-design-64dbbf12de82

Thakur, A. 2019. *The Lean Six Sigma Dictionary: Explaining the Basics of Continuous Improvement.* Delhi: ARTH-Excel.

Thermo Fisher Scientific. 2019. *User Interface Specification.* Thermo Fisher Scientific's internal material.

Thermo Fisher Scientific. 2020. *Technical Specification.* Brochure.

Turner, N. 2018. *6 key UX metrics to focus on.* [Accessed 26 October 2020].
Available at: http://www.uxforthemasses.com/ux-metrics/

UX team. 2020. Thermo Fisher Scientific. Group discussions March and April 2020.

Verma, N. & Sharma, V. 2020. *Energy Value Stream Mapping.* Bilaspur: Booksclinic.

Wiio, A. 2004. *Käyttäjäystävällisen sovelluksen suunnittelu.* Helsinki: Edita Publishing.

Zibell, K. 2017. *Choose the Right UX Metrics to Show Business Impact.* [Accessed 27 October 2020].
Available at: https://medium.com/@kristin.zibell/choose-the-right-ux-metrics-to-show-business-impact-7aafc5f33beb

APPENDIX 1: LEAN UX CANVAS – UX METRICS

Lean UX Canvas (v2)

Title of initiative: UX metrics

Iteration: 01

Business Problem

Currently the success of usability engineering and user experience design are not measured as efficiently as they could. Without proper measuring it is basically impossible to prove that the UX activities performed have been worthwhile and effective. How could the success be properly measured to help realize where to concentrate the limited UX design resources in use?

Solutions

Find ways to measure the ease of use and user experience.

Business Outcomes

The usefulness of UX work can be clearly justified and guidelines are created to guide where to concentrate in the design work.

Users

End users (laboratory analysts and laboratory technicians) who work at the laboratories and use the laboratory automation devices on a daily basis.

User Outcomes & Benefits

The end users spend less time in problem solving and more time doing the intended sample handling.

Hypotheses

I believe that the usefulness of UX work can be clearly justified if the end users spend less time in problem solving when the device is easier to use.

What's the most important thing we need to learn first?

How to measure the success of usability and user experience work?

What's the least amount of work we need to do to learn the next most important thing?

Find examples how the success of usability and user experience work is measured in other fields than laboratory automation devices.

APPENDIX 2: USABILITY TESTS

Background information about the test participants

	Job title	Laboratory experience in years	Familiarity with laboratory automation
User A	QC technician	1	Uses analyzers daily. Not familiar with laboratory automation.
User B	Laboratory scientist	10	Experienced with analyzers but has used also laboratory automation.
User C	Laboratory technician	7	Uses analyzers daily. Familiar with laboratory automation.
User D	Clinical research coordinator	5	Has worked in laboratories where laboratory automation is in use. Familiar with the systems.
User E	Laboratory manager	20	Has used on a daily basis.

Test cases and their success criteria

Test case ID	Test case description	Success criteria
1	Start the laboratory automation system.	All devices are powered on and in running mode.
2	Insert sample tubes into the system.	Sample tubes are inserted to the system successfully through the input device.
3	Stop the input device.	Input device is stopped successfully.
4	Solve the problem cases with the input device.	The problems are solved and the input device is running.
5	Perform the weekly maintenance actions for the input device.	The weekly maintenance actions are performed according to instructions.

6	Remove a carrier from the input de-vice.	Carrier is removed according to in-structions and the input device is run-ning.

APPENDIX 3: END USER INTERVIEW NOTES

User B

<table>
<tr><td>1. What kind of laboratory devices do you use in your daily work?</td><td colspan="2">Clinical chemistry analyzers.</td></tr>
<tr><td>2. How satisfied are you with the devices you use?</td><td colspan="2">The devices are very different. Some require more learning than others. Some devices guide you to perform the needed tasks.</td></tr>
<tr><td>3. What characteristics define if the device is good/bad?</td><td colspan="2">A good device has a simple user interface and guides the user to perform the needed tasks. Everything needed can be easily found. A good device is also quiet.</td></tr>
<tr><td>4. What are the characteristics that you value in laboratory devices?</td><td colspan="2">Ease of use and reliability. It is annoying if you need to constantly look for the information from the user manual.</td></tr>
<tr><td>5. What are the five most important characteristics from the following?</td><td><u>Ease of use</u>
Effectiveness
<u>Safety</u>
Reliability
<u>Scarcity of errors</u>
Visual appearance
<u>Quality</u>
Modernity</td><td>Speed
Simplicity
<u>Time saving</u>
Recommended by colleague
Reputation of the manufacturer
Number of features</td></tr>
<tr><td>6. What single characteristic do you think is the most important one?</td><td colspan="2">Ease of use and safety</td></tr>
</table>

User C

1. What kind of laboratory devices do you use in your daily work?	Liquid chromatography–mass spectrometry instruments. Instruments for blood analysis.	
2. How satisfied are you with the devices you use?	In general almost all devices have certain weaknesses. The weaknesses are mainly in the software user interfaces.	
3. What characteristics define if the device is good/bad?	Good: simple, easy to use, performing the intended purpose is simple, there are extra features for more advanced users, and one can easily find all the needed functions and features. Bad: graphical user interface is messy and unclear, performing the needed tasks is not possible even when using instructions.	
4. What are the characteristics that you value in laboratory devices?	Ease of use and clarity. The device must communicate its error state clearly and preferably even cope from the errors by itself. The devices produce a lot of waste so reducing the amount of it would be good.	
5. What are the five most important characteristics from the following?	**Ease of use** **Effectiveness** **Safety** **Reliability** **Scarcity of errors** Visual appearance Quality Modernity	Speed Simplicity Time saving Recommended by colleague Reputation of the manufacturer Number of features
6. What single characteristic do you think is the most important one?	Safety	

User D

<table>
<tr><td>1. What kind of laboratory devices do you use in your daily work?</td><td colspan="2">Different kind of analyzers: clinical chemistry, blood gas, blood count, hemoglobin device.</td></tr>
<tr><td>2. How satisfied are you with the devices you use?</td><td colspan="2">All of them have their own problems. I would expect same use logic from all devices from the same manufacturer. Laboratories anyway have devices from multiple manufacturers which makes it more difficult.</td></tr>
<tr><td>3. What characteristics define if the device is good/bad?</td><td colspan="2">Bad: analyzer does not recognize the need for dilution, devices do not work with certain samples, there is a lot of manual work (dilution, aliquoting samples from tubes to cups, checking the amount of sample in the container).
Good: device can tell if an error has happened.</td></tr>
<tr><td>4. What are the characteristics that you value in laboratory devices?</td><td colspan="2">Simplicity and ease of use. The user interface must be simple. It should be rather fast to learn how to use the device but it is ok that some learning is required. Maintenance of the device should also be straightforward but not happen too often.</td></tr>
<tr><td>5. What are the five most important characteristics from the following?</td><td>**Ease of use**
Effectiveness
Safety
Reliability
Scarcity of errors
Visual appearance
Quality
Modernity</td><td>Speed
Simplicity
Time saving
Recommended by colleague
Reputation of the manufacturer
Number of features</td></tr>
<tr><td>6. What single characteristic do you think is the most important one?</td><td colspan="2">Reliability</td></tr>
</table>

User E

1. What kind of laboratory devices do you use in your daily work?	I am familiar with laboratory automation systems in clinical laboratory and blood bank use.	
2. How satisfied are you with the devices you use?	The people working in laboratories always want better and more usable devices. Biggest problems happen if the automation system is not running or what needs to be done if there is a wrong result.	
3. What characteristics define if the device is good/bad?	Good: the use in general must be in accepted level. This means that the amount of stops and downtime caused by them are low. It should be possible to do maintenance for one device without disturbing the use of the rest of the system.	
4. What are the characteristics that you value in laboratory devices?	Ease of use. No long trainings required. User maintenance tasks are simple. More controls of the features should be given to the lab staff. Laboratories need a separate place to monitor the functioning of the automation system.	
5. What are the five most important characteristics from the following?	**Ease of use** Effectiveness Safety Reliability **Scarcity of errors** Visual appearance **Quality** Modernity	**Speed (Capacity)** Simplicity Time saving Recommended by colleague Reputation of the manufacturer **Number of features**
6. What single characteristic do you think is the most important one?	Uptime and modularity	

APPENDIX 4: PRODUCT MANAGEMENT INTERVIEW NOTES (HIDDEN)

APPENDIX 5: TENDER REQUIREMENTS (HIDDEN)

APPENDIX 6: A3 REPORT – UX METRICS (A3 DIVIDED INTO TWO A4 PAGES)

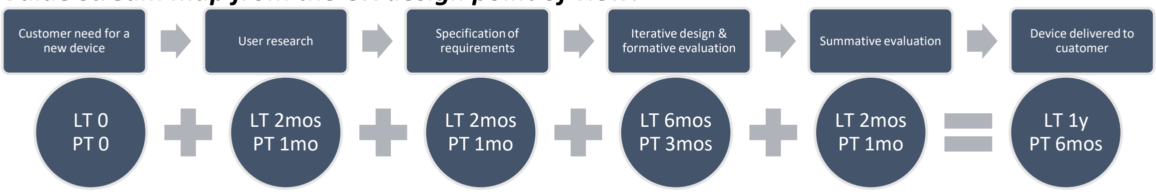
1. Problem space (from customer's perspective)- DEFINE

- ***Environment:*** *Product development of laboratory automation devices.*
- ***Problem:*** *The outcome and success of usability engineering and user experience design are not measured. UX design resources are limited and it is not possible or even sensible to invest resources in everything.*
- ***Critical-to-quality (CTQ) requirements:***
    - *A functioning device that fulfills its intended purpose*
        - *Use safety*
            - *Functioning safety features & no hazardous use situations*
        - *High uptime*
            - *Lack of errors & reliable device*
        - *Use without problems*
            - *Usability & user experience*
- ***Areas that need development:*** *UX metrics and resource allocation.*
- ***Objective:*** *To create UX metrics that help allocating the UX resources in the future.*

2. What is the process and the measurements that are used like? - MEASURE

- ***SIPOC (suppliers, inputs, process, outputs, customers) diagram****:*

- ***Value stream map from the UX design point of view:***



- ***What measurements are used? – Do they reflect the quality experienced by the customers?***
    - *Safety – the devices must fulfill certain safety standards. Safety reflects the quality experienced by the customers.*
    - *Effectiveness – the devices must function according to their intended purpose. Effectiveness reflects the quality experienced by the customers.*
- ***What new measurements should be taken into use?***
    - *UX metrics to measure the outcome and success of UX design.*

3. Root cause analysis – ANALYSE

- *Potential causes of the problem and root cause analysis performed using five whys method.*
    - *The success of usability engineering and UX design is only measured with use safety and use effec-tiveness.*
        - *It is not mandatory to measure other factors than safety and effectiveness according to IEC 62366-1 (2015).*
        - *To measure usability and UX in other ways requires planning.*
        - *It is unclear what to measure and what is a good enough level.*
        - *It is not easy to measure usability and UX.*
        - *Usability and UX consist of many features.*
    - *Resources are limited.*
        - *It feels that there are not enough personnel and time.*
        - *The resources in use are not fully utilized.*
        - *It is not completely clear where to concentrate in the user experience design.*
        - *There are no metrics that measure the success of the user experience design.*
        - *It is not known what the metrics should be.*
        - *Customers have not been asked from.*

4. Collecting and evaluating the proposed solutions - IMPROVE

	Little added value to customers	*A lot of added value to customers*
Easy to implement	• *Study what UX metrics are used in general.* • *Ask internal product management what qualities they value.*	• *Ask the end users what qualities they value.*
Time consuming to implement	• *Organize focus groups to find out from end users what qualities are valued.*	• *Organize usability tests to test UX metrics.* • *Ask the paying customers what qualities they value.* • *Create a set of suitable UX metrics for laboratory automation product development.* • *Test UX metrics in a full length project.*

5. Implementation plan for the design of experiments - IMPROVE

DOE	*Notes*	*Predicted results*	*Results*
Study what UX metrics are used in general.	*The preliminary list of UX metrics is presented in chapter 5.1.2.*	*Suitable metrics for laboratory automation devices can be found to work with.*	*Behavioral: Task success rate, time on task, search vs. navigation, user error rate.* *Attitudinal: SUS, NPS, CSAT.*
Organize usability tests to test UX metrics.	*Organized at the same time with the end user interviews.*	*Task success rate, time on task, search vs. navigation, user error rate all suitable.*	*Task success rate and user error rate turned out to be the most useful behavioral UX metrics.*
Ask the end users what qualities they value.	*Interviews held at the same time with the usability tests.*	*Qualities: ease of use, colleague recommendations, satisfaction.*	*Qualities: ease of use, safety, scarcity of errors, reliability.*
Ask internal product management what qualities they value.	-	*Qualities: ease of use, safety, scarcity of errors, reliability.*	*Qualities: uptime, modularity, reliability, modernity, visual appearance.*
Ask the paying customers what qualities they value.	*Product management interviews revealed lists of tender requirements from customers. They were used to replace the interviews of the paying customers.*	-	*Qualities: uptime, ease of use of the user features.*
Create a set of suitable UX metrics for laboratory automation product development.	-	*Behavioral: Task success rate, time on task, search vs. navigation, user error rate.* *Attitudinal: SUS, NPS, CSAT.*	*Behavioral: Task success rate (effectiveness), user error rate, safety.* *Attitudinal: SUS, CSAT, aesthetics, UUA, safety.*

6. Ensuring the solutions work - CONTROL

- *Suitable UX metrics for laboratory automation device product development were successfully created based on metrics used in other fields of industries, usability testing, and interviews with various stakeholders. Based on the end user and product management interviews and trying out some of the metrics in the usability testing, it can be assumed that the selected metrics work as intended.*
- *The measures taken have succeeded in creating metrics for usability engineering and user experience design and thus the objective initially created can be considered as achieved.*

7. Follow up - CONTROL

- *Update the usability engineering process to contain the new UX metrics that support the design work.*
- *Update appropriate document templates to include the new UX metrics.*