CHRISTIAN MARIAN

The development of a usability evaluation tool, integrating different factors that impact UX

São Paulo
2017

CHRISTIAN MARIAN

The development of a usability evaluation tool, integrating different factors that impact UX

Trabalho de Formatura apresentado à Escola Politécnica da Universidade de São Paulo para a obtenção do Diploma de Engenharia de Produção

Orientador: Prof[a]. Dr[a]. Uiara Bandineli Montedo

São Paulo
2017

FICHA CATALOGRÁFICA

AKNOWLEDGEMENTS

First, I would like to thank my family who supported me through my whole studies. They made it possible for me to have the best conditions to develop my studies and reach the best results possible. Without my family's support, these three awesome semesters abroad would not have been possible.

Besides that, I would like to thank both my universities, TUM in Germany and USP in Brazil, that gave me the opportunity to do a Double Degree program. I have learnt a lot for my further life in Brazil, I have met many valuable people and I have had an unforgettable experience in my 18-month exchange in São Paulo. I appreciate all exchange students and class mates who helped me to realize the tests for this research. Moreover, a special thank you to iPoli who did a wonderful job settling me into my new life in São Paulo, my tennis team at Poli and all friends I have made here in São Paulo.

Last but not least, I would like to thank Prof. Uiara Bandineli Montedo. She did a great job guiding me through my whole thesis. I appreciate the time and meetings she spent with me to help me and to reach the best possible result. Of course, I would also like to thank all other professors that guided me through my stay here at USP and who were part of my ideas that formed this thesis.

RESUMO

Hoje em dias empresas são conectadas fortemente pela internet no mundo digital. E-commerce, e-learning, entertainment e reservas online ganham mais importância na sociedade. Empresas tem que considerar o usuário e desenvolver interfaces favoráveis e que trazem satisfação ao cliente.

Além disso, os usuários se diferenciam bastante entre si. Desta forma, paginas da web que apontam para diferentes grupos da sociedade tem que considerar as diferenças do usuário. Algumas culturas e países teriam diferentes experiências se comparado com outros. O comportamento e a percepção do homem poderiam se diferenciar do comportamento da mulher. A afinidade do uso da internet mudaria com a maturidade. Todos esses aspetos têm de ser considerados pelo arquiteto do website quando a pagina da web apontar para vários grupos na sociedade.

A contribuição desse trabalho fica na conexão da usabilidade com fatores que poderiam impactar o uso da pagina da web. Além disso, essa pesquisa desenvolverá um método para avaliar a usabilidade a que está aplicado por uma pagina de web exemplaria. Essa avaliação será realizada por cobaias. Pelos resultados dos testes, tanto recomendações pela pagina de web serão feitas quanto diferenças entre os grupos diferentes são observadas.

Palavras Chaves: Usabilidade. User Experience. Método de avaliação. Cultura. Género. Idade.

ABSTRACT

Nowadays many businesses are connected to the internet and digital world. E-commerce, e-learning, entertainment, online booking are gaining more and more importance in society. Companies must meet the users' needs and must provide user friendly and pleasurable interfaces to their clients.

In addition to that, users differ in some ways. Especially websites that try to target as many user groups as possible, must consider these differences. Some countries may have more experience with the internet than others. Men may behave differently during their internet usage than women. Internet affinity may change with age. All these factors must be considered by the website designer when targeting many different groups.

The contribution lies in connecting usability with factors that may impact a website's usage. Besides this, the thesis develops a usability evaluation tool that is applied to an exemplary website by realizing a study with testers. According to the results of the tests, recommendations to the websites are given, as well as differences between certain test groups being observed.

Keywords: Usability. User Experience. Evaluation method. Culture. Gender. Age.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABREVIATIONS AND ACRONYMS

ENIAC	Electronic Numerical Integrator and Computer
GERD	Gender Extended Research and Development
HCI	Human-Computer-Interaction
HCU	Human Centered Usability
IoT	Internet of Things
LA	Latin American
OTA	Online Travel Agency
SUS	System Usability Scale
TUM	Technical University of Munich
UI	User Interface
USP	Universidade de São Paulo
UX	User Experience
WWW	World Wide Web

TABLE OF CONTENT

1 INTRODUCTION

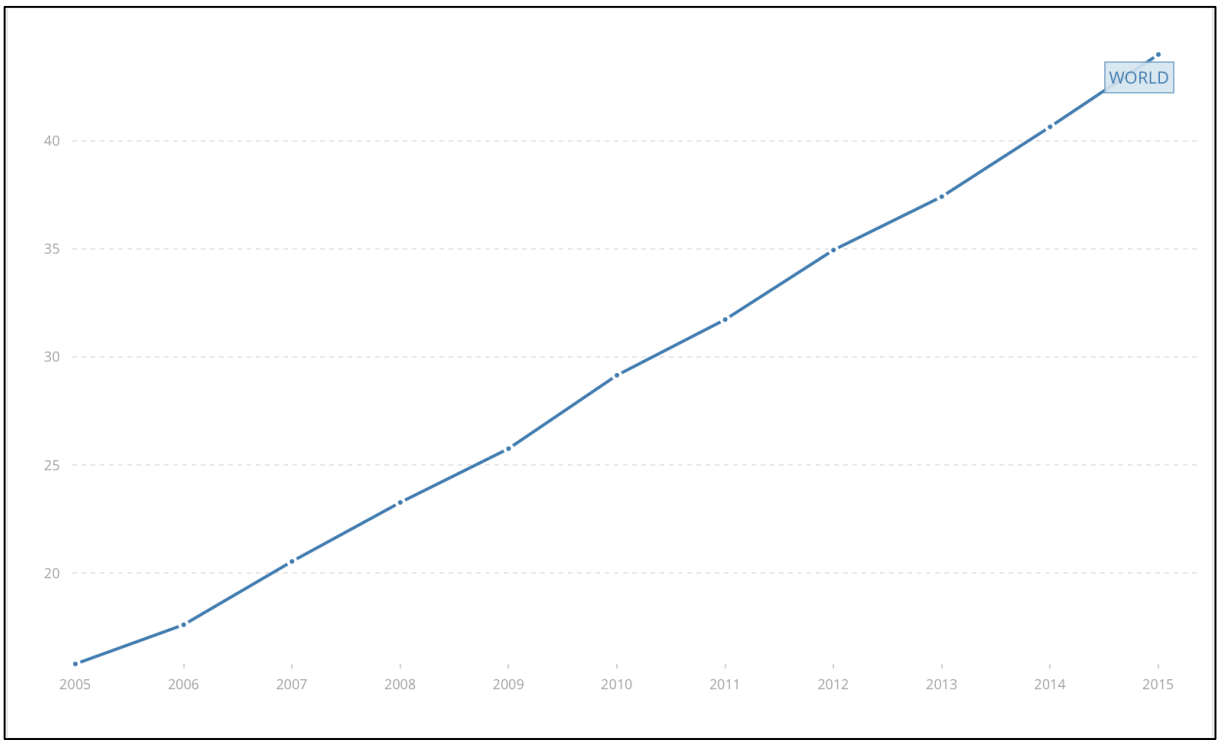
1.1 CONTEXT

The digital world and electronic devices play a more and more important role nowadays. Applications, e-business, university materials, online meetings and online calls are increasingly supported and improved by the internet.

Internet usage in general has increased massively in the past 10 years. In 2005, 15.8% of the population worldwide used the internet frequently. 10 years later, in 2015, more than 44% of the world's population used the internet frequently. (THE WORLD BANK, 2016)

This is an increase of almost of 200%. Figure 1 shows the worldwide trend of internet usage.

Figure 1 – Worldwide internet usage trend between (THE WORLD BANK, 2016)



Brazil's internet development is likely having the same trend. In 2005, 21% of Brazil's population used the internet. Nowadays (2015) almost 60% uses internet frequently, which is also an almost 200% increase. (THE WORLD BANK, 2016)

Germany is rather different on this topic. In 2005 more than 68.5% of Germany's population already used the internet, more than Brazil few years ago! In 2015, 87.6% of Germans were using the internet in their daily life. (THE WORLD BANK, 2016)

However, this comparison also shows that not every country can be treated identically regarding user background. On the one hand, in some countries (in this example Germany) the internet is more like a daily feature and is part of almost everybody's daily life. On the other hand, in other countries like Brazil or especially countries in Africa or Asia, the internet is not routine in daily life. According to The World Bank (2016) the average internet usage in Africa is merely 17.8% and Asia's average use is 35.7%. Excluding the exceptions (see table 1), which are basically far over 50% of the population's internet usage and which are well connected to the western countries, the usage decreases to 23.4%.

Table 1 – Outsiders of Asian countries (adopted to THE WORLD BANK, 2016)

Country Name	**2015 [%]**
Hong Kong SAR, China	84.9
Japan	93.3
Korea, Rep.	89.9
Macao SAR, China	77.6
Malaysia	71.1
Singapore	82.1

In the example above, cultural influences have been shown. However not only culture impacts the user's experience. Other factors, which will be discussed in this work, are essential for user success with internet platforms and electronical devices, as well.

1.2 PROBLEM DEFINITION

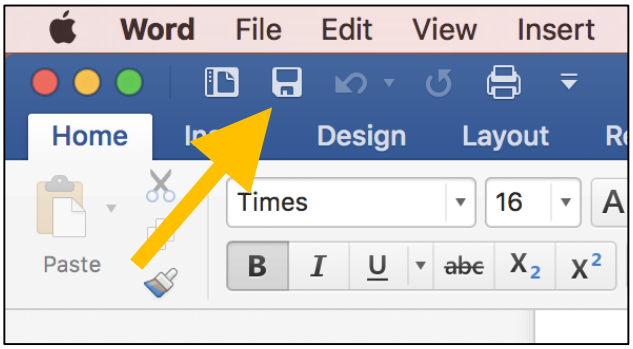
Before going on to the problem definition, the area must be defined. There are several types of technology or internet usage. Wendy A. Rogers and Fisk (2010) define it as "task being performed". This means there is a big difference if the device is supposed to be for professional use, like medical surgeries, or if it is supposed to be for everyone's use. Experts that use those

devices in their business are trained and are supposed to be used to these devices. However, if the technology is intended to be for many society groups it cannot be assumed that there is any knowledge regarding their profession. This research focuses on everybody's use which gives a wider target and bigger challenge for web designers. An example of everybody's use are hotel booking websites. A very wide range of users deal with these websites since many people use them to travel. It may be on business or on holiday.

As shown in the example above, not every country can be treated identically. Some people are used to internet usage; others are new in this domain and have little experience with the internet and electronic devices and others have not experienced it at all. Each type of user background (experienced, less experienced and no experience) has a different impact on the usage. Simple activities (e.g. the saving icon on the upper left corner on figure 2 or scrolling up and down with two fingers) might be easy and intuitive for some people but for others it might be not intuitive at all and would create a big challenge.

Interesting is that the saving icon symbolizes a disc which used to be common in early computer ages. Nowadays there are no saving discs anymore and newer generations have not even seen a disc like this. However, every mature or young person, who is familiar with computer use, knows that this button means "saving".

Figure 2 – Save button on upper left corner



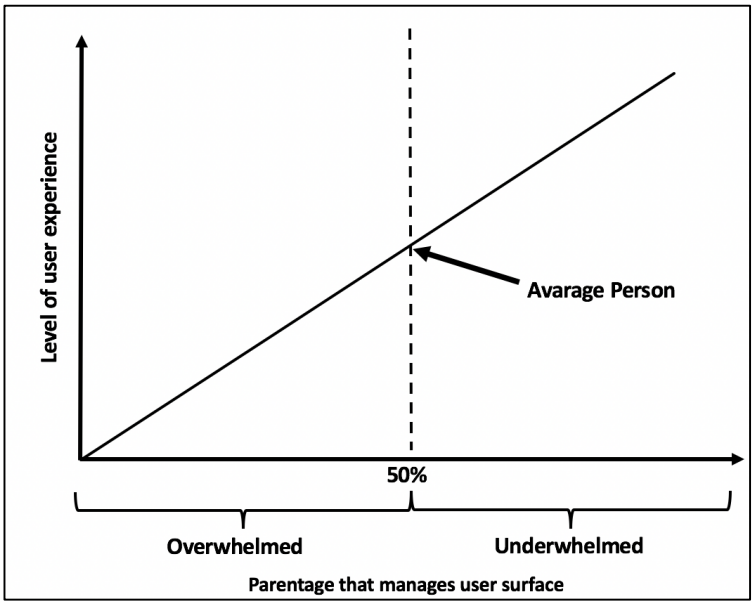
Many theories have been developed to evaluate a website's usability and handiness. However, most of those evaluation methods do not consider user background which influences the user's usage. Surveys, checklists, web designers or usability experts only focus on general rules and norms which were created for general usage.

Almost every expert agrees with the statement that the user must be in the center when designing electronic interfaces:

“Before starting to design a product, the user profile has to be completely clear. Motivation, requirements and level of qualification have to be defined.” (BAUMANN AND LANZ, 1998)

The same authors also mention that defining an average user is a big mistake because if only one type of person (the average person) is considered, the majority would be unsatisfied by the product because the range is way bigger than only the average person. Figure 3 shows the distribution, assuming the level of education is proportional (education is not proportional but in this example, it should help to illustrate the issue of the average person). This means the average person would be right in the middle of this distribution at 50%. If only the average person was considered, only this single person would have a comfortable and perfect usage of the user interface. The other half would be either overwhelmed or underwhelmed. This is the reason why the average person cannot be considered in general.

Figure 3 – Average person issue (adopted to BAUMANN AND LANZ, 1998)



Even if every web designer thinks he puts the user in the center, he often does not. The user’s background is not considered deeply, which means not every essential factor that influences the user background is considered. The most common questions to identify the user are (HALE, 2014):

- Which are the user's requirements?
- What does the user want?
- Who are the main users?

These questions are important, indeed. But these questions should be broken down into more specific question because if not, the risk of ignoring specific groups and generalizing the user is huge. Questions to avoid this problem can be, for example:

- Where is the user from?
- What kind of educational background does he have?
- How old is the user?

It is kind of understandable that those questions are mostly ignored because this takes time and makes the problem and the web designing more complex. However, this work should give an overview of the most important factors which should be considered when designing and evaluating user interface.

1.3 OBJECTIVES

The main goal of this work is to create a usability evaluation method to evaluate website and to identify weak points of the website considering different user factors, such as age, culture or gender. To reach this main objective, the research questions should be listed:

- How to develop a usability evaluation tool considering different user factors?
- Which factors may influence User Experience?
- How much impact do these factor have on the user's usage?

To reach a satisficing answer of the questions above, an extensive overview of existing usability evaluation methods and of already existing factors that influence User Experience should be done in the further steps. Additionally, tests that deal with the objective should be executed to discover the impact of these factors.

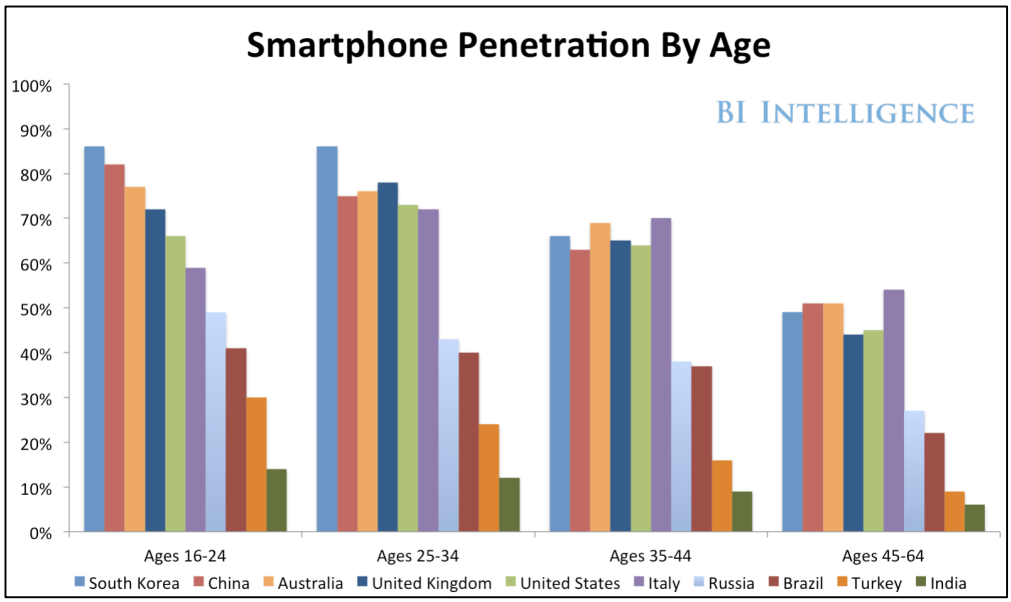
1.4 RELEVENCE OF THE WORK

Business via internet is getting more and more important. E-Shopping like Amazon, Mercado livre or EBay are getting bigger and bigger. Physical stores like supermarkets (e.g. Pão de Açúcar, Carrefour, Walmart, REWE) or shops (e.g. Galeria Kaufhof, Apple, Zara, H&M) are working closer and closer with the internet. They publish promotions, new collections or even sell their products online. Travelling (booking flights or hotels) is nowadays an internet business. According to Forbes (2015), 75% book their trip online.

Since internet has had such a big influence on today's business and it will get even more important, the user interface must be attractive and the user must enjoy using it and should not end up in anger or frustration, which is unfortunately often the case. (HALE, 2014)

As already shown in chapter 1.1 culture may impact someone's user experience. On the one hand, there are countries where children grow up with smartphones and usually have their own smartphones at an early age. In 2013, according to Cocotas (2013), around 80% of South Koreans, Chinese or Australians between 16 and 24 years old owned a smartphone. In contrast, less than 40% of Indians, Turkish or Brazilians had a smartphone. Consequently, some countries experience the digital life earlier than others. Figure 4 shows the smartphone penetration by age for different countries.

Figure 4 – Smartphone penetration by age (COCOTAS, 2013)



Web designers tend to focus on their biggest markets and on people who use the internet frequently, because they generally bring the most income. However, looking at this from a different perspective and studying the statistics, there is more potential in the less developed internet market than in the mature internet market like in Germany. So, this group should not be ignored, in contrast, they must be considered carefully and intensively.

Germany's internet usage grew from 2005 to 2015 from 68.5% to 87.5%, which means an increase of 27.74%. Brazil's internet usage in 2005 was 21%. In 2015, 60% of Brazil's population used the internet, which is an increase of 181%. The potential in less developed internet countries is even bigger:

In 2005 Asia's population used 12.44% the internet (excluding exceptions it was: 3.23%). In 2015, already 35.7% used it (excluding exceptions: 23.23%). This means an increase of 187% and, excluding the exceptions, an increase of 619.2%. There has also been a big increase in African countries. In 2005 only 2.84% people used the internet frequently and in 2015, it increased to 17.8%. This means a relative increase of 526.76%.

Table 2 – Countries mentioned with their increasing rate (adopted to THE WORLD BANK, 2016)

Country Name	2005 (%)	2015 (%)	Increase (%)
Asia (excluding outliners)	3.23	23.4	624.5
Africa	2.84	17.8	526.8
Asia (including outliners)	12.44	35.7	187.0
Brazil	21.02	59.08	181.0
Germany	68.7	87.6	27.5

Analyzing the table above, the relative trend is more interesting for less developed internet countries (Asian or African countries). In other words, in the years ahead many people with low internet experience will enter the digital market. For this reason, these people are very important new potential customers.

Nevertheless, the cultural aspect is only one of many essential factors that should be considered. Other factors like age groups, gender or educational level also influence user experience and usability.

1.5 OUTLINE OF THE THESIS

The approach of this thesis is divided into 10 chapters.

In the following chapter 2 a theoretical basis will be given. In 2.1 important definitions, will be explained. Cognitive Ergonomic, which is part of the user's behavior will be explained in 2.2 which is followed by the difference between Usability, UX and UI.

In chapter 3 different literature is presented that show, on the one hand, various methods to evaluate usability. On the other hand, various authors presenting existing factors that impact user background. At the end of the chapter, the deficiencies of the discussed literature are pointed out and the focus of this thesis is clarified.

In chapter 4 the most important factors that impact user experience are defined. Some factors have been worked out in chapter 3 and others will be introduced. At the end of this chapter, it will be illustrated that considering user experience factors is not easy and that each user experience factor influences another one.

Having all essential factors that impact the user's background, a usability evaluation method to measure the importance of each user experience factor will be created in the following chapter 6. Chapter 5 will explain the steps taken to develop the method and will give some explanation of the following evaluation method.

As soon as the evaluation method has been created, a software based tool will be developed to support it. This tool will be presented in detail in chapter 7. Followed by an application of this tool in chapter 8 which evaluates a hotel online booking website and compares different user experience factors.

Finally, chapter 9 and 10 will show the results and conclusions drawn from the exemplary application of the tool in the previous chapter.

2 THEORETICAL BASIS

This chapter should give a clear overview of Usability, User Experience and User Interface. First, chapter 2.1 will define key words which will be important for the further research. Cognitive ergonomics is an important scientific area regarding this topic. A brief explanation will be given in chapter 2.2 In chapter 2.3 the differences between Usability, User Experience and User Interface will be highlighted. Even though there is a connection between those expressions it is important to understand the differences before going deeper into this research.

2.1 DEFINITIONS

In the following, key words of this research will be explained.

Usability

According to DIN EN ISO 9241-11 (1998), Usability is the effectiveness, efficiency and satisfaction with which specified users achieve specified goals in particular environments.

User Experience

User Experience (UX) refers to a person's various experiences when interacting with a product (Kim, Han et al. 2015). In this work the product will be specified by a website or an application.

User Interaction

User Interface (UI) are all components of an interactive system (software or hardware) that provide information and controls for the user to accomplish specific tasks with the interactive system (ISO 9241-100 2010).

Human Computer Interaction

Human Computer Interaction (HCI) is the study of the way in which computer technology influences human work and activities. The term ‘‘computer technology’’ nowadays includes most technology from obvious computers with screens and keyboards to mobile phones, household appliances, in-car navigation systems and even embedded sensors and actuators such as automatic lighting. (LIU AND ÖZSU, 2009)

Internet

The Internet is a worldwide system of computer networks. It gives the opportunity to talk to users at other computers, as well as share information in case users from other computers give their permission. (ROUSE, 2005)

World Wide Web

The World Wide Web (WWW) is not a synonym for the Internet. While the internet is a network of network where all information resides, the WWW makes it possible to access to pages by using a Web browser. (CHRISTENSSON, 2006)

Website

A website is a file that relates to the World Wide Web. The address of this file belongs to an individual or to a company. The beginning file is the home page which allows to get to all other pages on this site. (ROUSE, 2005)

2.2 COGNITIVE ERGONOMIC

Human cognition searches, treats, stores and uses different types of information of the environment and creates knowledge which will be included in future actions of the human being. In other words, the process starts with the memory that compares known situation to the new problem, analyzing and choosing the best solution which leads to decision making. (ABRAHÃO, SZNELWAR, et al. 2009)

In this context perception and cognition play an important role which cannot be considered differently whereas it must be considered as a continuing process. The perception takes place when an object is observed by its characteristics and realized by the human sense (color, size, format etc.). Cognition is involved as soon as the object is interpreted by the human being and processed by memories, categorizations or problem solving. (ABRAHÃO, SZNELWAR et al., 2009)

The ergonomic cognition deals with the cognition during the work rather than with the function of the human cognition. That means if a worker brings up information and experience from former work or other areas and applies them to a new task, he successfully used its

cognition. The individual does not always have all information and knowledge about the situation. To avoid bad results, in these cases, the goal of ergonomic cognition is to build systems as intuitive as possible to manage systems even if not all information is available. (ABRAHÃO, SZNELWAR et al., 2009)

The discussed ergonomic cognition shows the importance of intuitive products. If some artefacts distract the user's work, the result can lead to accidents or to an unsatisfied result (ABRAHÃO, SZNELWAR et al., 2009). Abrahão, Sznelwar et al. (2009) also mention that different people come up with different cognition which can be helpful to combine them to get successful result for complex problems (shared cognition). Anyway, if a person cannot count on the cognition or support of others because the person is simply alone at that moment, it is important to reduce the UX to the desired minimum to satisfy most target users possible. To avoid any kind of distracting artefacts, it is important to consider usability to build successful and intuitive UI.

2.3 USABILITY, UI AND UX

After defining important words for the further thesis, the differences between Usability, UX and UI will be clarified. Even though these words are deeply connected to each other, there are essential differences.

Difference between Usability and UI

UI is basically the series of screens, pages and visual elements like buttons or icons that a user manages to interact with a device. The very first computers did not have icons or graphics. The computer looked more like its own computer programming language. In the 90's Apple started a revolution and designed the Macintosh which was the beginning of user-friendly interfaces. It was the first step of commercial success of a computer with multiple windows and point-and-click mouse. (LANOUE, 2016)

According to Nielsen (2014), usability is a quality attribute of UI. Usability deals with the question: Can the user accomplish the goal? And UI builds the interface that deals with the question of being able to accomplish the goal (NIELSEN, 2014). So, UI makes Usability primarily meaningful. And Usability is an important challenge in making UI successful.

Difference between Usability and UX

The first requirement for an exemplary UX is to meet the exact needs of the user without bothering him. Simplicity and elegance are the second. The user must enjoy using the product. (Norman and Nielsen 2016)

According to one of the most recognized UX experts, Donald Norman (2016), everything influences the user experiences. Even the experience in the shop when the user buys a product or the product's transportation from the shop to their homes is part of the UX.

It is important to know much more than only what the user wants. To achieve high quality UX, website designers must consider multiple disciplines, including engineering, graphical and industrial design and interface design. (NORMAN AND NIELSEN, 2016)

As already mentioned, usability is a quality attribute of UI. According to the definition of Usability, websites must be easy to use, must be efficient and should satisfy the user. While usability literally only focuses on managing the physical interface, UX goes way deeper. It is not only the experience with the interface it is also everything else which influences the user's experience and satisfaction.

Difference between UI and UX

Many experts have given many examples regarding the difference between UI and UX. UX pioneers Norman and Nielsen (2016) describe it as following:

> It's important to distinguish the total user experience from the user interface (UI), even though the UI is obviously an extremely important part of the design. As an example, consider a website with movie reviews. Even if the UI for finding a film is perfect, the UX will be poor for a user who wants information about a small independent release if the underlying database only contains movies from the major studios. (NORMAN AND NIELSEN, 2016)

Lanoue (2016) uses Google as an example. He praises Google for having an awesome UI. On the Google website, only the logo, search bar and a few buttons are placed – simple and everything necessary appearing. However, users have access to almost the whole human knowledge in a few seconds. This satisfies the user, who enjoys using it again and again. Now

imagine it always takes more than 15 seconds for each research – even if the UI remains – the UX would change dramatically.

Ken Norton, partner at Google Venture, compares the difference between UX and UI with a restaurant. UI is the table, chair, plate, glass and utensils. UX is everything from the food, to the service, parking, lighting and music. (LANOUE, 2016)

2.4 SUMMARY

This chapter has given an overview of the important theoretical basis which should simplify the understanding of this research. Definitions of important words have been given. Ergonomic cognition has been introduced and shown that it is important to consider the user's cognition, as well. The main differences between the key words Usability, UX and UI have been clarified.

In the next chapter usability evaluation methods and literature that point out different factors that impact usability and UX will be discussed. This should demonstrate the need of this research and the lack of science.

3 STATE OF ART

This chapter gives an overview of literature that has already dealt with the research questions mentioned in chapter 1.3 on page 25. Because the research questions address, on the one hand, usability evaluation methods and, on the other hand, UX factors, this chapter will be divided into two parts which give examples of each main research area. In chapter 3.1 the most important existing usability evaluation methods are worked out and in chapter 3.2, literature that has already dealt with UX factors is presented. Finally, in chapter 3.3 the lack of science will be underlined.

3.1 EXISTING METHODS

There are basically four ways to evaluate user interfaces. There are formal methods using some analysis techniques, automatic methods using computerized procedures, empirical methods which are experiments with testing users and heuristic methods which look at the interface and pass judgement according to one's own opinion. (NIELSEN AND MOLICH, 1990)

First, a formal evaluation method is described which takes advantage of technology such as eye tracking, mouse clicks and other analytical techniques. Then questionnaires are described which is a very important empirical usability test. After that a heuristic evaluation is described, followed by usability criteria and hints by Bastien and Scapin (1993). And finally, an evaluation method, by the University of São Paulo, that takes advantage of empirical, formal and Bastian and Scapin criteria is presented.

A Comparison of Eye Movements When Searching Web Pages between Americans and Koreans

Rau (2013) describes a study, done by Valdosta State University, GA, USA, that compares American students' behavior looking at random websites to Korean students. They were asked to look at 30 different international websites for 10 seconds. Their eyes were tracked by an eye tracking system and the students were asked, after looking at the website, about their experience and comfort while doing so. Different criteria were defined such as fixation, count on each area of interest or time spent on each fixation area.

The conclusion of the test was that Koreans' fixation time was longer than the Americans' and Americans might pay more attention to banner images while Koreans paid more attention to viewing navigation areas.

Questionnaires

A very common method to evaluate usability are questionnaires. Users are asked to fill out surveys, afterwards the surveys are analyzed and problems may be identified. Surveys are part of qualitative evaluation methods. This means the human being is in the center of the evaluation and opinions, feelings and observations are considered, whereas quantitative methods produce a objective and quantitative (measurable and numerical) result. (U.S. DEPARTMENT OF HEALTH & HUMAN SERVICE, 2016).

Most surveys use a Likert scale which gives the option of answering a question between a defined range (mostly between "I completely agree" and "I completely disagree"). One well known and common usability evaluation survey is the System Usability Scale (SUS) by John Brooke (1996). John Brooke (1996) developed a method to evaluate usability systems cheaply and quickly (in technical language: quick and dirty). After a deep theoretical description of usability, he concludes that evaluating usability quantitatively is not effective at all and questionnaires are strongly recommended.

John Brooke developed a questionnaire with different statements that should be rated according to their validities. These statements ask for the feelings and the opinions during the usage of the system. The validity is rated by a Likert scale which ranges from 1 to 7. One means the statement does not match at all with the user's opinion or perception. Seven means the user agrees completely with the statement. John Brooke's SUS questionnaire can be found in appendix A.

Heuristic Evaluation of User Interfaces

Nielsen and Molich (1990) created 9 principles of heuristic evaluation of user interfaces. Even though heuristic evaluations are difficult, they may have a big importance. For example, if time and lack of expertise do not allow for an empirical method (experiments), heuristic methods

are highly recommended. Other advantages are the cheap and easily planned execution and usage in the early stage of the project.

A heuristic evaluation is basically looking at the interface, discovering the weak points of the interface and improving it. It is usually done by the designer and does not need any experiments. The above mentioned 9 principles of heuristic evaluation are:

1. Simple and natural dialogue
2. Speak the user's language
3. Minimize user memory load
4. Be consistent
5. Provide feedback
6. Provide clearly marked exits
7. Provide shortcuts
8. Good error messages
9. Prevent errors

Criteria of Bastien and Scapin

The French ergonomists Bastien and Scapin (1993) developed 8 usability criteria to design and evaluate HCI. Some criteria have sub-criteria which extend the list to a total of 18 criteria (see table 3). Above all, the criteria should be a supplement of other usability test methods rather than one evaluation method itself. Considering only the criteria can be very analytic and may miss important evaluating points.

Table 3 – Usability criteria (Adopted by BASTIEN and SCAPIN, 1993)

Nr.	Criteria	Sub-Criteria	Sub-Sub-Criteria
1	Guidance	Prompting	
		Grouping/Distinction of Items	Grouping/Distinction of Location
			Grouping/Distinction of Format
		Immediate Feedback	
		Legibility	
2	Workload	Brevity	Concision
			Minimal Actions
		Information Density	
3	Explicit Control	Explicit User Action	
		User Control	
4	Adaptability	Flexibility	
		User Experience	
5	Error Management	Error Protection	
		Error Correction	
		Quality of Error Message	
6	Consistency		
7	Significance of Codes		
8	Compatibility		

Ergonomia e Usabilidade: Em Ambiente Virtual de Aprendizagem

The book written by Abrahão, Montedo et al. (2011), professors of the *Universidade de São Paulo* (USP), describes a method that improved a website (in this case the Moodle website) which was regarded difficult to use by many professors at USP. To improve it, they analyzed the current situation. There were frustrations and verbal unhappiness, which can partly be considered as a think aloud method.

After proving the need to improve the website, a quantitative evaluation of the Moodle tool was done. The use of this tool was evaluated by criteria such as errors that occurred during the usage.

Different professors were given tasks to carry out with this tool (e.g. upload course materials on Moodle). Each professor's solving procedure was documented by registering the number of clicks, mouse movement, recording their action or task success and navigation.

After analyzing the results of the task, the improvement of the Moodle interface started by considering the criteria of Bastien and Scapin. In total, three iterations have been done to come to the result.

3.2 EXISTING FACTORS

To connect the importance of literature to already existing UX factors, table 4 was created. The left-hand side shows the literature that is related to the research. The top right side shows the last two research questions from chapter 1.3 on page 25. In between, the circles show how much the literature deals with each research question. In the following the different meaning of the circles are explained.

- ● - The literature covers a lot of the research question
- ◑ - The literature covers part of the research question
- ○ - The literature does not cover anything of the research question above

Table 4 – Literature and research questions

Research Questions / Literature	Which factors influence UX?	How much impact do these factors have on the user's usage?
Patrick Rau	●	◑
Yrjö Engeström	◑	◑
Frandsen-Thorlacius, Harnbaek et. Al.	◑	○
Geert Hofstede	◑	○
Corinna Bath	◑	◑
Rogers and Fisk	◑	●
Janine Liebal and Markus Exner	◑	◑
Marsden and Kempf	◑	○
Donald Norman	◑	●
Thüring and Mahlke	●	◑
Johannes Robier	◑	○

A short summary of the illustrated literature in table 4 will follow. The literature that deals extensively with the research question will be explained in more detail after a short summary of all mentioned literature.

Patrick Rau

The Cross Culture HCI International Conference in 2013 brought about many interesting articles about culture and web designing. Comparisons between different countries like friend-making on social networks, user evaluation of websites in Arabic society or comparison of Australians and Taiwanese websites. (RAU, 2013)

Because almost each topic of this conference is interesting regarding showing the culture influence on digital usage, usability and UX, some articles will be discussed in the further work.

Yrjö Engeström

The Danish professor Yrjö Engeström took the Russians Activity Theory and applied it to educational technology tools. The activity theory became an interdisciplinary framework which also gained importance in the HCI (CLEMMENSEN, KAPTELININ et al., 2016). In one of Engeström's pieces of literature, he deals with advancing virtual reality, 3D and mobile technology which apparently matters when considering usability, UI and UX (MWANZA, ENGESTRÖM, 2005). Later, one piece of publication that describes educational tools including the activity theory will be discussed in detail.

Frandsen-Thorlacius, Harnbaek et. Al.

The study, Non-Universal Usability? A Survey of How Usability is Understood by Chinese and Danish Users, from Frandsen-Thorlacius, Hornbæk et al. (2009), compares a survey filled out by Danish and Chinese people. This survey was filled out by 412 participants (154 Danish and 258 Chinese) and the answers showed that the Chinese and the Danish have different interpretations of successful usability. The Danish put their focus on effectiveness and no frustration while the Chinese place importance on emphasis on visual appearance, satisfaction and fun.

Geert Hofstede

Professor Geert Hofstede (2017) created a cultural model that consists of five dimensions. Each dimension group is found in a society that was empirically discovered and occurs in combination with the other dimensions. The five dimensions in Hofstede's model are power-distance, collectivism / individualism, uncertainty avoidance, masculinity, and long-term orientation. Because lots of literature refers to Hofstede's model, his model will be described more deeply at the end of this chapter.

Corinna Bath

Corinna Barth's (2009) doctor thesis, *De-Gendering informatischer Artefakte. Grundlagen einer kritisch-feministischen Technikgestaltung,* (English: Basics of critical feminist technical design) deals with the question of how can software be de-gendered. She lists different theories of gender differences or discrimination and in the following she gives techniques on how to avoid

gendering. One reason for gendering is the "I-methodology" which means that designers completely ignore different user knowledge and skills. They consider themselves as the user and do not consider people with less UX, less computer skills, less money etc. To avoid this error, she recommends the persona-method which tries to consider a wide range of different user types.

Rogers and Fisk

Toward a Psychological Science of Advanced Technology Design for Older Adults, by Rogers and Fisk (2010), proves that relatively, the use of the internet is increasing among the older age groups rather than the younger age group. Consequently, they call for focus on improving the usability considering older people. In the end, they discuss which areas should be focused on and what differs between an older person and a young person regarding their internet usage.

Janine Liebal and Markus Exner

Liebal's and Exner's (2011) book *Usability für Kids* (English: Usability for kids) deals with usability for teenagers and kids. They explain the behavior and interests of kids in their different age ranges. They show the importance of focusing on this age group and they define the different activities that children engage in online. After showing the differences between adults and children ("Children are not adults") and different activities between girls and boys, they point out different practical rules which should be considered when designing for kids.

Marsden and Kempf

In Marsden's and Kempf's (2014) book, Gender UseIT, different gender theories regarding technology are given. Examples of jobs, where men have better chances to climb the hierarchical ladder or IT jobs that have also been done by women, were given. A method to understand the user (female or male) will be presented as well as differences between analyzing websites. Since gender will be one factor in chapter 4, this literature will be summarized more deeply.

Donald A. Norman

Donald A. Norman (2002), one of the UX and usability pioneers, explains in his book, The Design of Everyday Things, good and intuitive design. He starts with examples of designing successful and unsuccessful simple objects like doors or cameras, and he advances to the design

of digital intuition. Since Donald A. Norman has been a very influential and important person in UX, his classic book will be discussed later in this chapter.

Thüring and Mahlke

Thüring and Mahlke (2007) put in practice various experiments regarding emotions, aesthetics and usability in their study Usability, Aesthetics and Emotions in Human–Technology Interaction. They compared attractive devices with unattractive devices, as well as well-designed and badly-designed interfaces. They measured subjective feelings like user emotions or happiness by measuring the heart rate, analyzed pupil responses or letting the participants fill out surveys. Their results were that UX is divided in two characteristics. On the one hand, instrumental qualities, such as effectiveness of its functionality or controllability of the system, are indeed important. However, on the other hand, UX depends on non-instrumental qualities such as visual aesthetics or haptic quality.

Johannes Robier

The book *Das einfach und emotionale Kauferlebnis* (English: The Simple and Emotional Buying Experience) by Johannes Robier (2015), deals with emotions during buying or dealing with something. It discusses how a product can influence somebody's thinking. Surprises, animals, personalization, sexuality, fear or humor can result in positive emotions. They explain the key factors of making a product interesting, which are simplicity and emotion. If a person can understand the situation directly and the person feels personally targeted, it influences the product positively. Consequently, design can bring the product and customer together by emotions. They also point out that a personal memory will be improved by activating more than one sense. If the person can listen to and see the product at the same time, the human's memory keeps the product in mind better than using only one of these two senses. If the person can be active, it increases the memory even more.

In the following part the most important and interesting literature mentioned above is examined in more detail.

Patrick Rau – Cross Culture Design

As already mentioned in the beginning of this sub-chapter this literature brings many different studies which compare different cultures to each other. In the following, some papers that deal with culture usability are discussed.

The study Incorporating Culture in Website Design: A Comparison of Taiwanese and Australian Website Characteristics shows once again that two different countries (Taiwan and Australia) do matter in many different points in UX. The study refers to experts that have had a lot of experience in web designing, computer science and international culture. They analyzed different Australian and Taiwanese websites. Their analysis was divided into five categories: visual presentation, navigation, links, layout and multimedia. Consequently, the websites were evaluated quantitatively regarding the experts' categories (e.g. did the home page have a search feature? Yes or no?). After all analysis, some differences between the two cultures have been found. Taiwanese sites usually count on moving pictures, stream videos, flash animation while Australian websites incorporate stream video and flash animation. Consequently, the Taiwanese prefer multimedia and moving pictures on their websites whereas Australians prefer static pictures and text. Moreover, the Taiwanese prefer using pictures and videos of charismatic and popular persons whereas Australians usually use pictures of daily life. (RAU, 2013)

The paper Developing a Contextual Network for Indigenous Communities in Mexico explains how to make indigenous people from Mexico familiar with unknown technologies. Communities in Mexico were visited and indigenous people were shown and asked to use devices they had not seen in their lives before. Usually these communities do know cell phones, televisions and radio but e.g. iPads were unknown. The purpose of this project was to bring indigenous communities together with new technology and after the team left, they were able to continue using the technology due to successful and intuitive usability. A challenge was the written language because many indigenous people could not read and some did not speak Castilian. However, by using different methods and symbols, indigenous people experienced fun by using the devices rather than frustration.

Anyway, this study proves that good usability and User Centered Design (UCD) allows all people with poor UX to use technology successfully. According to this, to realize perfect UCD it is important to consider the following points (RAU, 2013):

- perfectly understand users and their tasks
- development of empirical measures of product usage
- the use of an iterative design

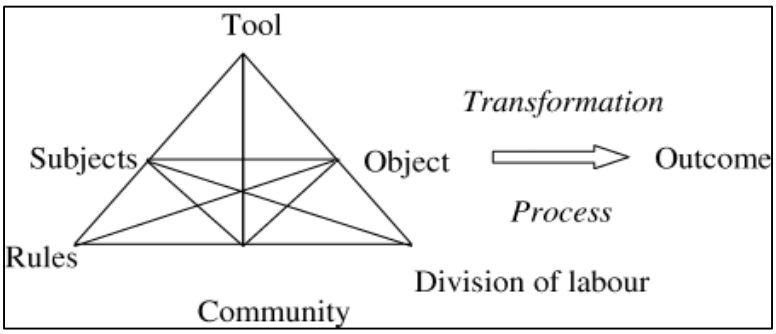
Yrjö Engeström

The activity theory was originally created by Russian psychologists Vygotsky and Leontiev. Yrjö Engeström, Miettinen et al. (1999), first dealt with this theory outside the USSD. On the one hand, he added the element of community to the Russians theory and on the other hand he also connected the activity theory to HCI. (CLEMMENSEN, KAPTELININ et al., 2016)

Mwanza and Engeström (2005) applies the activity theory on his work to consider learning theories to the task of managing content in e-learning. Basically, the activity theory system consists of different components that drive an activity. This leads to the activity triangle model. The activity triangle consists of subjects that interact with objects to achieve a desired outcome. This means that human interaction is mediated to each other and elements like tools, rules and division of labor influence the activity. Mediators represent the nature of relationships that exist within and between participants of an activity in each community of practices. Figure 5 shows the activity triangle model.

Engeström claims that e-learning does not address social-cultural and pedagogical aspects of users. Mwanza and Engeström (2005) compared paper based learning forms to digital learning forms in participating schools. Different teaching scenarios and environments were tested and the conclusion of the interaction between people, organizations, technologies and information that impact the learning objects and educational content, was given.

Figure 5 – Activity triangle model (MWANZA, ENGESTRÖM, 2005)



Geert Hofstede – Cultural Dimensions

Professor Geert Hofstede did one of the most recognized studies about culture and country differences (HOFSTEDE, 2017). He collaborated with IBM between 1967 and 1973. By analyzing vast databanks, he created a comparison between various countries. In the beginning, he started to compare 40 most influencing countries and expanded till reaching 76 countries that could be compared to each other. He came up with five dimensions which characterize the behavior of the people of each country. In the following part the different dimension are defined by Hofstede (2017):

Power Distance describes the extent that the lower classes of a society accept the fact of unequal distribution of power and how this inequality is handled within a society. A large Power Distance suggests a society where these inequalities are simply accepted, whereas a low Power Distance shows a society where people demand equality and do not accept unjustified inequalities.

Individualism: Individualism defines a society where people take care only of themselves and their immediate family. In contrast, the opposite is collectivism, whereby society is closer and people look after their relatives and other relevant groups and receive firm loyalty in return.

Masculinity: A society represented by Masculinity favors a more competitive kind of society where achievement, materialism and success are strived for and rewarded. A society represented by Femininity, on the other hand, favors quality of life, cooperation and is more caring. These two contrasting societies are sometimes referred to as “tough versus tender” cultures.

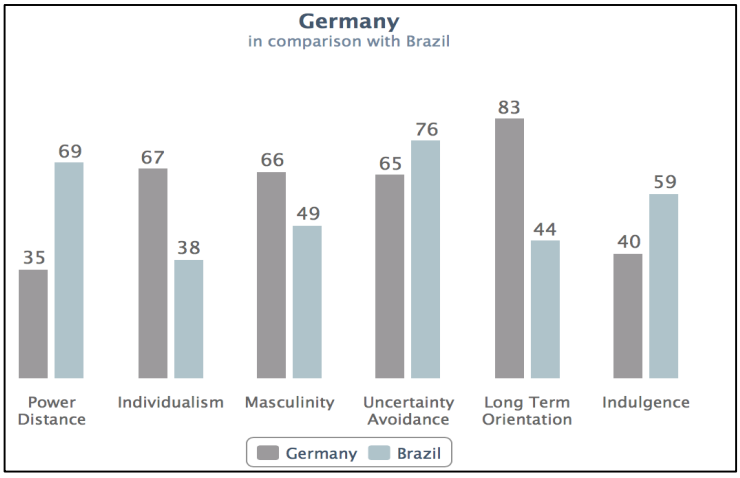
Uncertainty Avoidance: The Uncertainty Avoidance (UAI) refers to the extent to which a society is able to live comfortably with uncertainty. Basically, how well does a society deal with the fact that the future will never be known. Countries with strong UAI have strict rules and dislike more new or unusual behavior. Those with weaker UAI are more relaxed.

Long Term Orientation: Societies who score low on Long Term Orientation are more traditional and dislike changes. Those with high scores are more practical and encourage education as a way to embrace the future.

Indulgence: An Indulgent society allows for the pursuit of fun and enjoyment of life, whereas as a society displaying restraint imposes strict behavioral norms and suppression.

On Hofstede's internet website two or more countries can be compared to each dimension. An example of Germany versus Brazil can be found in figure 6.

Figure 6 – Comparisons Germany versus Brazil by Hofstede's 5 dimensions (HOFSTEDE, 2017)



Rogers and Fisk – Toward a Psychological Science of Advanced Technology Design for Older Adults

Rogers and Fisk (2010) discussed the aging problem of technology. Older people have more difficulty than younger people. This age group should be considered and technology made pleasurable for everybody's use (both older and younger people). According to this study, older people use fewer technology devices than younger adults do. In a survey with 17 common

technologies, older adults reportedly used an average of 12 of those technologies. Compared to the younger adults, who used on average 15 out of the 17 technologies.

This study also proved that the increase in technology use is mostly happening among the older age groups, those in their 70's and older. The internet use between 2005 and 2009 increased, for example, from 26% to 45% in this age group.

Anyway, the stereotype that older people do not like technology is disproved by this study. 62% of older people who have been interviewed do like to use the internet and only 38% would rather keep their distance form it. So, this shows the importance of considering different age groups and their different UX.

Marsden and Kempf – Gender UseIT

Differences in gender are still political, cultural and business issues. Studies prove that female and male technology users still have different technology skills, even in developed countries like Germany. This literature consists of different points of view regarding gender differences in IT, Usability and UX. Some authors agree with gendering and others disagree.

One theory says that gender is not born. Gender depends on doing, attitude and behavior. Gender can even change during lives. This phenomenon is called Doing Gender. Consequently, two assumptions can be defined:

1. Gender distribution changes/power and role of women changes over time (e.g. 15 years ago it was unbelievable that women would occupy leading roles in politics. Nowadays women do influence politics. An example is the female German chancellor, Angela Merkel)
2. Gender cannot be considered as Women or Men (Dual-view). It is more a dynamic process. If there was only "man" or "woman", social status, sexual interests and physical conditions would be ignored.

Furthermore, this book also presents some aspects that influence usability and UX. Gardenswartz and Rowe's dimension theory is presented. Three dimensions show that gender, age, culture, job, education and technical background have an impact on usability and UX. Gender is especially important for websites with pictures because those are considered differently by men and women. Age is important because different health conditions worsen for older people

(older people suffer from bad eyesight, have movement restriction etc.). Considering hobbies is important because the expectation varies regarding the activities in their free time. Students may have different expectations than married couples. And education, for example, influences the understanding of language. Moreover, women consider some applications more important than men. That means women spend more time on those applications than men, which affects the UX as well.

Donald A. Norman – The Design of Everyday Things

This classic book by Donald Norman (2002) claims that the simplest and most obvious things in daily life can lead to frustration because they are badly designed.

Donald A. Norman questions new, wonderful, revolutionary and good devices. He questions why do good and new devices need instructions or consultants who must explain how these devices work? This book explains how product-design can fail and how to avoid misunderstandings and frustrations while using products. There is no way to avoid human error but there is a way to minimize it by good design. The key for creating a successful product is usability and aesthetics. In this book, many examples of successful and unsuccessful products are given. When talking about products in this book, not only digital devices are considered. One examples that shows that users are extremely important to the product's success is shown in the following:

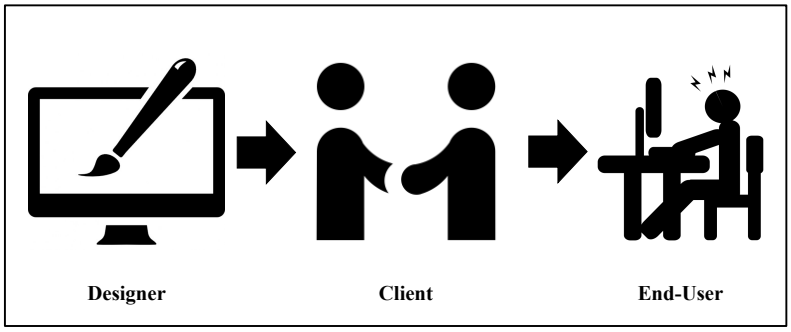
Norman gives an example of two buildings that were designed differently. One building was designed and planned only by architects. The other building respected employee's wishes and concerns. Consequently, the employees of this building were more satisfied and improved their efficiency by 7% while the efficiency of the other building was unchanged.

Furthermore, the main problem of failing is called "folk psychology" or "naive psychology". Designer are not typical users and this is the reason why their designs fail. Designers know their products and have the proper knowledge – the user often not.

Another problem is that designers often confuse users with clients. Designers mostly design for clients who will deliver for the users. This means designer create their product for the client

and not for the end user. The client often thinks about costs. Consequently, the design is focused on reducing costs which means usability and users suffer. Figure 7 illustrates this issue.

Figure 7 – Designer works for client who publishes for end-user (adopted to NORMAN, 2002)



Norman (2002) concludes that new products and new technology do not only improve someone's quality of life. They can be extremely confusing and can lead to frustration if they are badly designed. Therefore, he strongly recommends following the user-centered design which puts the user in the center of designing.

3.3 LACK OF SCIENCE

A successful website requires a good user interface and good usability. This was discovered long time ago and has been studied for decades. Evaluating websites regarding their usability is a common procedure and improvement tips are common as well. Furthermore, the discovery of the diversity of users is nothing new either. Moreover, the importance of dealing with different users is discussed at length in various literature. Theories and experiments that describe different behaviors of cultures and age groups or even genders are studied in great detail.

Only a minority of literature connects different factors and considers them when designing websites, even though a few studies discuss an evaluation method that consider UX factors. Most evaluation methods address an overall user or a specific user group for their specific system. However, they forget that users can have the same interests but come from different countries, are in different age ranges or even vary in their gender.

Considering table 4, on page 39, most literature does not satisfy the first research question (Which factors influence UX?). Most of it only considers one single UX factor but one person is influenced by many different factors. Even though Rau (2013) and Thüring and Mahlke (2007)

do mention multiple UX factors, they are poor at connecting them to the question how much they impact the user's behavior.

Hofstede (2017) compares countries nicely, its culture behavior and evaluates different dimensions. This obviously means there are differences between countries and cultres but he does not apply this to usability and user interfaces. In addition, he only considers culture and does not evaluate other factors like age, emotion etc.

Again, Marsden and Kempf (2014), Rogers and Fisk (2010) and Norman (2002), each of them worked out interesting factors that impact UX. However, these studies are separately done and only consider one single UX factor.

This thesis should intervene at this point and put together the most important factors that impact usability. After finding and listing the most important UX factors, an evaluation method should be developed which evaluates websites and measures the importance of each UX factor.

3.4 SUMMARY

This chapter gave an overview of scientists and literature that has already dealt with the topic of this research. First, the most essential usability evaluation methods were shown.

In the next sub-chapter, different studies on existing UX factors were presented. While listing them, it was realized that most of the factors were considered in isolation and were not put in context to other factors or to their impact on usability.

In the next chapter, the most important UX factors that may impact usability are summarized, partly based on the literature just discussed and other factors newly introduced.

4 UX FACTORS

This chapter summarizes, on the one hand, the most important factors that have been mentioned in the previous chapter and on the other hand, completely new UX factors will be introduced that may impact usability, as well.

The already mentioned factors include culture, age, gender, emotions and educational degree. The two new factors to be introduced to this research are, firstly, the impact of using the internet in someone's free time or at work and secondly, the current occupation. These newly defined factors are based on experience in relation to the economic and industrial field. The derivation of this situation is illustrated by a short case study in chapter 4.4. The second factor, Occupation in chapter 4.5., was discovered by life experience. The comparison between surfing on the internet without any rush, to doing so on a busy time schedule at work was noticed.

4.1 CULTURE

Culture is one of the most researched UX factors. Many scientists have researched the influence of culture on usability and UX and a lot has been written about it. Culture can be very complex which includes country of birth, country where living, religion and so on.

There are obvious differences between countries, people and their behavior. This was illustrated by the Dutch professor Geert Hofstede and other authors like Patrick Rau in chapter 3 (page 43 and 45). For instance, Rau (2013) proved that Australian and Taiwanese people care about different features on websites. While Taiwanese people care about multimedia and moving pictures like streaming videos or flashing animations, Australians prefer static pictures and texts on websites.

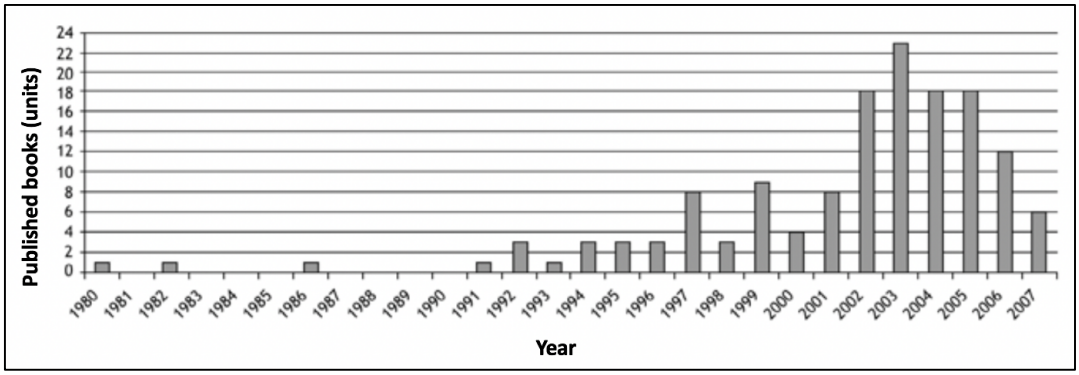
This research assumes that the reason for these differences might be the difference between cultural behaviors. People from different cultures grow up differently and behave differently. The access to the internet also differs from country to country as shown in statistics above. While some kids have grown up with smart devices and the internet, other kids from other countries have not experienced this in their childhood.

4.2 AGE

"Children are not adults" (LIEBAL and EXNER, 2011). This sentence is obvious and it is also obvious that children have different interests and behave differently. This is the reason why age is also an interesting and important UX factor that may influence usability.

Probably this factor is not discussed as much as culture but it is still topic of books and many articles. Liebal and Exner (2011) wrote about usability for kids and they published a statistic about the historical development of literature that dealt with age as an UX factor. Figure 8 shows the years between 1980 and 2007 and how much literature was published each year. The first official literature was published in 1980. However, in the 90's it became more popular and in 2003 it reached a peak. After this year, the published literature decreased again.

Figure 8 – Published literature on children's computer usage (LIEBAL and EXNER, 2011)



According to Liebal and Exner (2011), children see the world differently and have different interests than adults. In contrast to many adults, children are egocentric. They mostly think of themselves and cannot understand other points of view. This phenomenon changes in school. In addition, they obviously do not have the same experience and knowledge as adults, since adults have more life experience.

Other phenomena are the reaction time and the mouse clicking time. On the one hand, the body is not fully developed like an adult body. Muscles and the coordination are not as developed as adult muscles. Furthermore, children might not have seen the task before, while adults may have and can take advantage of their experience. (LIEBAL and EXNER, 2011)

So, children are not adults but seniors are neither children nor adults. It is obvious that older people take more time to understand new technology or they tend to get frustrated while using technology (ROGERS and FISK, 2010).

Older people usually use less technology than younger people. Again, reaction time is different and may take longer than a younger adult does. Changes in perception, movement control and cognitive capabilities also must be considered. For instance, the eye sight of seniors is worse which is the reason for the mentioned differences. (ROGERS and FISK, 2010)

4.3 GENDER

Anyway, Liebal and Exner (2011) argue that girls and boys care about different things especially in early ages. Boys usually prefer video games and girls usually prefer chatting. According to them, gender also differs in colors. Girls find different colors attractive than boys. That means the perception of websites might be different between girls and boy. The same authors also argue that boys usually start using computers earlier than girls which leads to having more experience with technology than girls (LIEBAL and EXNER, 2011).

Nevertheless, another theory claims that gender is not born. It changes over time and depends on education. For instance, 15 years ago it was unbelievable that women would occupy leading roles in politics or other job areas, nowadays, they do. (LIEBAL and EXNER, 2011)

There are jobs where women also have potential to climb the hierarchy ladder but others are still men focused. According to Bundeszentrale für politische Bildung (2010), the German center of political education, 40% of women have leading positions in health and social area professions in Germany. However, according to them, only 8 to 18% of women have a leading position in banks or in insurance companies.

Even though the gender role might have been changed in the recent yerars, there may be still differences regarding women and men.

4.4 EMOTIONS

Sluggish system response obviously leads to frustration but futuristic and innovative design can result into surprise, curiosity or pleasure which influence the use of the product positively. (NORMAN, 2002)

To avoid frustration, the user must be in the center of the design and the designer must be aware of this. In many cases the designer assumes himself as the user which can lead to trouble. The user usually does not have the same background and knowledge as the designer. The user may see the interface the first time while the designer has worked on it for a long time. If the designer assumes himself as the user, the whole system can lead to misunderstanding, non-intuitive use which results in inefficiency and lack of motivation. (NORMAN, 2002)

Kevin Hale (2014) also points out in his speech at Stanford University that the user must be the priority. It does not help if the interface is the most beautiful one but the user does not understand it. He underlines the fact that designers or web engineers must work directly with the user and completely understand the user. His start-up, for instance, works 4 to 6 hours per week only on customers' issues to improve the problem of the platform and make the whole platform user friendlier.

A great feedback is the SUS questionnaire by John Brooke (1996) that asks for the user's happiness and satisfaction after using the website.

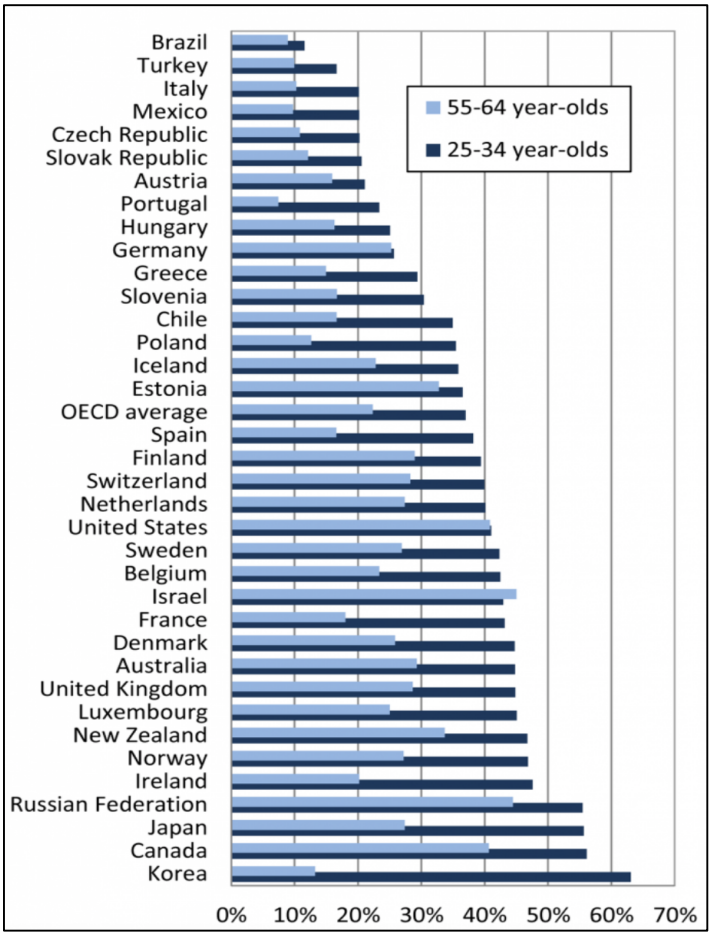
4.5 EDUCATIONAL LEVEL

Educational level may have a connection to the cultural factor because in developed countries a higher percentage of the population have a higher education than in developing countries. Nathan E. Bell (2011), director of Council of Graduate Schools in the USA, published an article that compares the attained tertiary education in different countries. Tertiary education includes, for example, higher education (BELL, 2011). He divides the population into two age groups (25-35-years-old and 55-65-years-old). In some countries, there are differences between these age groups and in others not. At this point it should not be considered deeply, but it should be mentioned that besides the country criterion, age also might influence the difference of educational level.

Figure 9 demonstrates the tertiary education difference regarding the statistics of Council of Graduate Schools by Bell (2011).

However, country and education should be considered separately because even in developing countries, there are many highly educated people. Considering Brazil, which is classed as a developing country, only 10% attained a tertiary education according to Bell (2011) but on the other hand there are also highly educated persons that have experienced high class education and universities.

Figure 9 – Percentage of population has attained tertiary education, by age group (BELL, 2011)



4.6 JOB VS. FREE TIME

There is obviously a difference if a website is used for business purposes or in someone's free time. To illustrate the differences between the internet usage during business and free time, a short case study will be presented:

> Mark started one year ago at a consulting company and is an analyst. His average working hours in the office are from 9.00 am to 7.00 pm (excluding traveling hours to clients). During his work, he uses for any analysis, the Microsoft operation system. For this purpose, Microsoft Office packages (Word, Excel and PowerPoint) are more favorable to use than any other operation system. Privately Mark is an iOS fanatic. He loves all Apple products and has an iPhone, iPad and MacBook which he uses for private purposes. One reason for his love for Apple is his former DJ career. In his college time, he sometimes played music as a DJ at parties and the iOS operation system is unbeatable for this purpose.

This case study used the two mentioned operation systems, Microsoft and iOS, just as an example. It could have been the other way around or two completely different operation systems.

Now it is interesting to ask the question what are the differences between using his digital devices during work and during his free time. To answer this question let's analyze Mark's daily life and situation. Some key points are:

- He has a long working day (over 10 hours)
- He uses Microsoft during work
- He prefers iOS in his free time
- He also uses only Apple products in his free time
- He loves Apple products and has good memories of using them

Considering these key points, following assumptions can be done:

- The relationship between his Microsoft computer and him is his job
- Due to the iOS cloud, he has all information and data automatically on all three devices which makes it comfortable to use and besides that, he creates a positive relationship due to his positive memories as a DJ
- After a hard-working day, he prefers to use his iOS products which take on the role of a familiar friend

In conclusion, after a hard-working day, normally a person wants to calm down and not think as much as at work, or at least do something that is pleasurable. This means the concentration is different between work and being at home on the couch after working 10 hours. A person wants to be in his comfort zone, not necessary think a lot or make a lot of effort. This

means the user interface should be even easier to manage, more pleasurable to use and intuitive, to create a comfortable and familiar atmosphere. At work, people's capacity and capability is usually higher than at night on the couch after a long working day.

4.7 OCCUPATION

Occupation might have a connection to the factor emotion in chapter 4.4. However, while occupation is rather an external factor that is driven by external conditions, emotions refer to the interface itself. The feeling and the perception during the usage are incorporated by emotions. The occupation depends on the person's background and current situation. In which situation is the person? Is he stressed by his job? Is he busy? Does he have many other things to do and is in a hurry? These are all external factors that influence the usage. The interface itself cannot change the person's situation but the interface can support the person by making the interface easy and quick to use, effectively and efficiently.

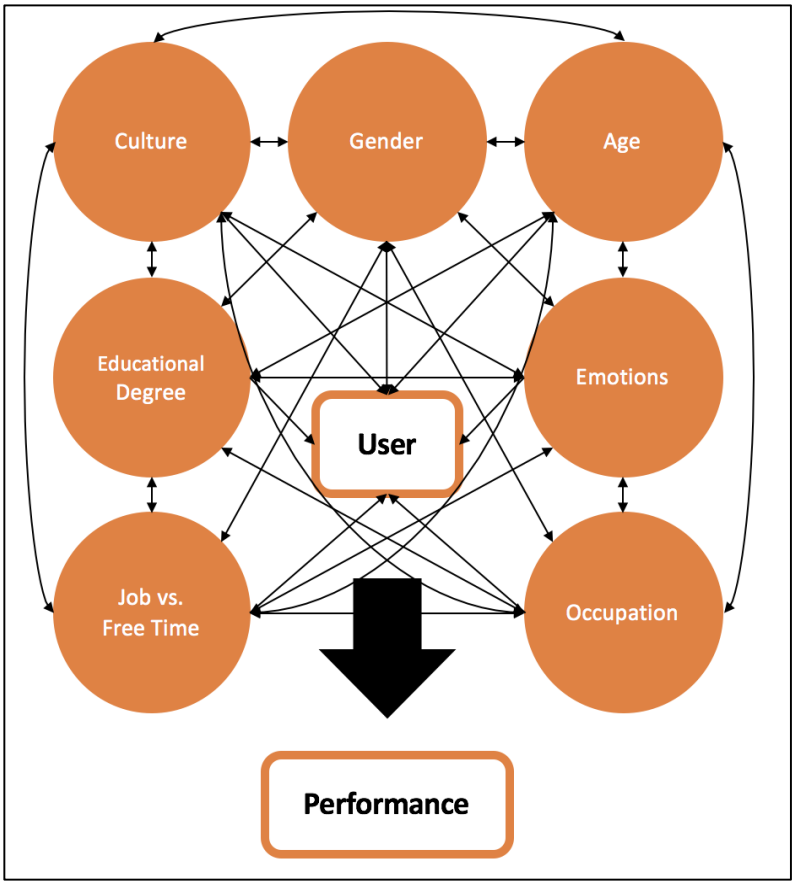
In the following an example from the industry is described: Amazon offers the One-Click-Buy. It basically helps the user to save time when buying products. After buying a product and after filling the personal information, such as name, address, credit card number etc., the user can buy other products within 15 minutes without filling in this information once again. The user only has to click on the product and he automatically purchases the product. This saves a lot of time if a user wants to buy different products. The purchase can be canceled within 15 minutes to avoid unwanted actions. (AMAZON, 2017)

4.8 UX FACTORS INFLUENCE UX

In the previous sub-chapters, different UX factors have been defined. Unfortunately, these factors cannot be considered separately. They influence each other. This means, for example, a 70-year-old German user cannot be considered as a 30-year-old German user, even though they may have the same nationality. However, age also plays an important role which makes the system more complex by each factor added because each factor correlates to the other factors in some way.

All in all, this study assumes that UX factors influence the user performance, the user's background but is also strictly connected to the other factors. Figure 10 below illustrates the complexity of integrating UX factors and the connection to each element.

Figure 10 – UX connected to other factors



4.9 SUMMARY

In this chapter seven factors have been presented that may impact user's usage. Besides that, this chapter has also illustrated the correlation between the factors themselves and the influence on the user as well as on the performance. The following chapters will discuss a usability evaluation method and how to measure the impact of those factors on the user's usage.

5 CLASSIFICATION AND DEVELOPMENT OF METHOD

This further research is divided in three main parts. The first part is the developing of the method. The second part is the application of this method which is deeply described in chapter 8. And the last part is the discussion of the results and drawing conclusions of them (chapter 9, 10).

The development of this method is also divided in three chapters (chapter 5, 6 and 7). The first chapter will show requirements that have to be met by the method, the website type that will be targeted and the core user that is interested in using this website type, as well as limitations on the discussed UX factors will be discussed, due to time and resource limitation. The second chapter will present the task that will be executed by the testers. In chapter 6.2 hypotheses about the testers' performances will be made which is followed by the data collection. This method will be measured quantitatively, qualitatively and by auto evaluations. How they will be measured exactly will be explained in this sub-chapter. In 6.4 the exact evaluation methodology will be explained. In the last sub-chapter of chapter 6, the execution process will be presented which was repeated for every tester.

After the development of the usability method a software-based tool, which is intended to simplify the execution of this evaluation method, will be presented in chapter 7. In chapter 8 an exemplary application of the method will be shown. The results of this application were analyzed and conclusions drawn in chapter 9 and 10.

5.1 REQUIREMENTS OF THE METHOD

This sub-chapter should give an overview about the requirements of the method itself. Most requirements are adopted by the doctor thesis of Irrenhauser (2014).

Usable

This method can be interesting for companies that focus on e-commerce or companies that use internet websites to reach customers. The company can evaluate their websites and check if their interfaces follow usability rules and consider UX factors. This method will offer an objective evaluation method to evaluate the productivity and efficiency of the interfaces. (IRRENHAUSER, 2014)

Practicable

This point requires the economic feasibility. The method should be easily enforceable, modest time consuming and should not require high financial resources. Adequate equipment and resources must be ensured as well. (IRRENHAUSER, 2014)

Flexible

The method should be flexible. This means the method should be useful for different purposes (IRRENHAUSER, 2014). This method especially needs to be useful for different websites with different content.

Transparent and Reproducible

The structural development should be comprehensive at all time (IRRENHAUSER, 2014). This leads to a serious and credible method. The experimental environment, the development and the results are clearly illustrated and can be reproduced by all interested scientists. To realize this point, it is important to explain the methodical development step for step.

Complementary

This requirement has not been mentioned in any literature but it is still important to consider it for this evaluation method. It should be possible to expand the method at any time. If new UX factors are discovered they should be able to be integrated in the method without any time-consuming effort or complex processes.

5.2 WEBSITE DIVISION

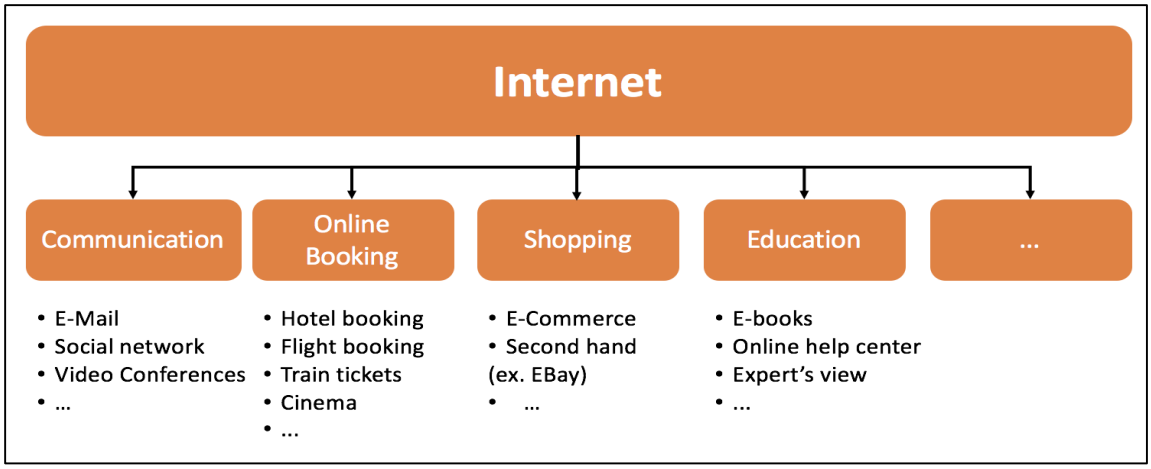
The internet is a huge business. Almost all information can be found on the internet. However, the internet also targets different people by different content and websites. To get an overview of online opportunities, the internet use will be divided into different areas.

Top 10 List (2013) published an article about the top 10 purposes of the internet. Communication is on the top position because it allows everybody to communicate from all around the world. Business, families, social media and many other target groups use the internet to communicate their interests. Online shopping is getting more and more popular, as well as

online booking. People do not like to wait in lines to be attended to. Nowadays, hotels, flights, cinema tickets etc. can be bought directly online. Figure 11 divides the internet into one of its most important user functions.

Due to the huge complexity of the internet, this study must choose one specific area which will be focused on. Since the travel business is very important and targets many different people, this study will focus on a hotel booking website.

Figure 11 – Internet divisions (adopted by TOP 10 LIST, 2013))



5.3 USER PROFILE

This method has two core goals. First the method should discover if the discussed UX factors in chapter 4 really influence the user's behavior. Second, this method should analyze a hotel booking website objectively. Is the tested website user-friendly and does it satisfy the core customer or is there any potential improvement of this website?

Before starting to develop any product, according to Baumann and Lanz (1998), the user must be clearly defined. As already discussed in the previous section, this research will analyze a hotel booking website. Following assumption are made about the online booking user which will be considered as a potential user:

- The user has access to the internet.
- The user can be either male or female.

- The user has access to the website at any time of the day. The user may be at work, at home or on the move.
- Every nationality can be considered, as in the globalized world almost every citizen of any country can travel.
- Income can vary as well. Some people travel low-budget and some people are looking for luxury travels.
- The user's education can vary massively. People with higher education may book a hotel but also people with basic education may take advantage of these websites.
- The considered age range will be 17 years and older. The research assumes that people under 17 years usually travel with their parents and the parents book the hotels. And since retired persons also like to travel, there is no limitation to the upper scale.
- The user can have very different background. They can have a lot of experience in booking a trip, since he has done it a lot but he can also have little experience. He can be a busy person who has limited time but he can also be a person who has a flexible day schedule.

5.4 UX FACTOR LIMITATION

Due to time and resource limitation, some UX factors cannot be considered at this research and others have to follow assumptions which makes it possible to evaluate them.

As shown above, culture can be very complex and is influenced by many factors, such country of birth or religion. Since time and resources are limited in this research, not every single factor that impacts culture can be considered. This research will limit culture on nationality and continents. It will assume that nationality and continents have a big impact on culture.

Age was already limited on people older than 17 years old (see user profile in chapter 5.3). Age will be simplified by defining different age ranges that may show representative results.

Gender will be limited on female and male. Changing gender, which was mentioned in the literature research, will not be considered.

Only a higher education will be considered in this research because the testing environment will be a university, assuming every tester is at least a student.

Emotion will be evaluated by the tester's opinion in form of a questionnaire. All other factors will not be considered in the further work because of its resource limitation. A person's occupation or stress level is difficult to measure, since occupation varies from person to person. To define a stressed person without any medical reports and medical consultation is almost impossible. Some people may be busy and may have many things to deal with but they do not consider themselves to be stressed while others may be stressed even if they have less to do. Moreover, the UX factor Job vs. Free Time will not be considered neither because it is impossible with this research's resources to create a working or free time environment during the tests.

5.5 SUMMARY

In this chapter, some general requirements on the method have been discussed. Starting with the requirements that must be met at all times by the evaluation method. Followed by a short division of the most important functions of the internet and by selecting the function that will be analyzed in this research. Assumptions about the possible user that usually books hotels online were made. Even though the target range of people that book hotels online is huge, chapter 5.3 limits the core user by listing different assumption of the user. In the end of this chapter, elimination and limitations of the considered UX factors have been done.

6 DEVELOPING A USABILITY EVALUATION METHOD

According to Kempthorne (1952), in scientific research there are two types of experiments, absolute and comparative experiments. Absolute experiments determine a specific characterization. And Kempthorne (1952) defines comparative experiments as following:

> A comparative experiment […] is an experiment in which two or more treatment are compared in their effects on a chosen characteristic of the population. In such an experiment the value of the characteristic under either treatment separately for the population is of no particular interest. (KEMPTHORNE, 1952)

Since different UX factors have been defined and the importance of each UX factor should be defined, a comparative experiment will be done. The goal is a conclusion about which UX factor may have influence on the user's performance and which may not. Moreover, the analysis should also discover the weak points of the website.

6.1 TASKS

The tester completed 2 different tasks which are presented in the following:

Task 1

The first task was completing a reservation on the hotel booking website Booking.com and cancelling it immediately. Before even starting the booking process, the user should change the website's language into a comfortable language (recommendable in the mother language or English). To challenge the user and to guarantee a fair evaluation environment, the booking process starts in Mandarin, assuming that no tester speaks Mandarin and is not comfortable at all with this language. Consequently, every tested person starts by understanding nothing and will switch to a more comfortable language immediately. The booking process will have the following constraints:

- Use a comfortable language to book a hotel
- Book a hotel in Rio de Janeiro, Brazil at Copacabana
- Booking dates: December 15th till December 17th

- 2 rooms with 4 persons in total
- Breakfast included
- The hotel must be cheaper than $1.000 in total (US Dollars)
- No credit card required for booking, no prepayment and free cancellation

As soon as the user has completed the reservation with the constraints listed above, he must cancel it immediately. When seeing the cancellation confirmation on the screen, task 1 was completed.

The destination Copacabana was chosen because it was assumed that every exchange student at USP is familiar with or has at least heard of Rio de Janeiro, Copacabana. This should ensure that every user had the same starting conditions and nobody had to deal with an unknown location.

Task 2

After the booking and cancellation process, the user should contact the customer service. The purpose of this task is to establish if it is easy to contact Booking.com. A random text was composed to be sent to the customer service. The message should be exactly as the following text:

> Dear Booking Customer Support,
>
> I apologize for cancelling my reservation but this is a university study and I didn't aim to travel to Rio de Janeiro, Copacabana on these dates. Anyway, as soon as I travel to Rio, I will book on your website and I intend to travel to the hotel I just cancelled.
>
> I am sorry for the inconvenience

A supporting sheet will be handed to every user doing this test. This has the advantage of giving user all constraints and both tasks in one overview. Appendix B shows the supporting sheet with all information mentioned above, that was handed to every tester.

6.2 HYPOTHESES

Without ideas or theories, the investigator is unable to choose which elements to observe (Kempthorne 1952). According to Kempthorne (1952), a hypothesis must provide an answer for a particular problem for which it was formulated. Moreover, it must give a prediction of observations that can be verified. He also claims that there are three classes of how to classify a hypothesis: Either the hypothesis falls in the class which accepts the hypothesis or it falls into the class where the hypothesis is rejected. And the third class is where the experiment is insufficient and the conclusion is that more experiments must be conducted to come to a reasonable result.

Since two tasks have been defined and UX factors are observed, the following hypotheses will be made which may or may not be verified in the further work. First the hypothesis will be formulated shortly, which is followed by a logical explanation of this hypothesis:

Booking.com is user-friendly and only few user (<5%) will fail on task 1

Booking.com is considered as a user-friendly website by many people. However, this website is not perfect and even when most users will complete the booking process successfully (task 1), few will fail on task 1 which will be defined by less than 5% failure.

More user will fail on task 2 but majority will complete it successfully

Since the core function of Booking.com is the hotel booking process, the second task (contacting the customer service) is a more specific task which might be a little more difficult for some users. Most users will manage the task but more users will fail compared to task 1.

Latin American will take more time than Europeans

Since a lot of literature claimed cultural differences regarding internet usage, at least in the booking process, differences between culture will be observable. The assumptions are based on Hofstede's 5 dimensions, particularly Individualism. Individualism means the extent to which a person/culture is dependent on others or the extent to which the person is used to work alone. Since European countries usually score higher on this point which means they are more used to work alone, Europeans will be quicker and will solve the task with more confidence.

Exchange students have more experience than Brazilians

Another hypothesis will classify exchange students with Brazilian students. Since exchange students only spend a limited time in Brazil, they want to discover the country and the continent and take advantage of weekends and holidays to travel around. In contrast, Brazilian students usually live with their parents and if they travel, in many cases the parents take care of the hotel. This is the reason why exchange students have more experience in online booking processes.

Women analyze the hotel details deeper than men

Some authors are convinced that there are differences between women and men (LIEBAL and EXNER, 2011). This hypothesis will claim that men solve the booking process in a straighter forward manner without paying attention to many details whereas women will pay more attention to details, beauty, location, reviews etc.

The older the person the more experience he has

This research assumes that most teenagers below 17 years used to travel with their parents and the parents usually took care of the hotel booking. Getting older this responsibility may change. The older a person the more he has booked hotels which, consequentially, means that this person has more experience.

6.3 DATA COLLECTION

The data collection is based on numerical and quantitative data and on auto evaluation or qualitative data. The numerical and quantitative data will be collected automatically by a supporting software during the execution of both tasks. The auto evaluation or qualitative data will be delivered by the user himself. He will answer a questionnaire at the beginning of the test, which asks basically questions about his booking experience. In addition to that, he will also answer a questionnaire after the execution which focuses on his satisfaction during the execution.

In the following, the qualitative data collection will be discussed in detail. Both questionnaires will be presented. In the second part, the quantitative data collection including the supporting software will be explained.

6.3.1 Qualitative Data Collection

Qualitative evaluation methods ask for someone's opinion, perception, feelings or experience (AUSTIN and SUTTON, 2014). An important qualitative method are questionnaires. In the following the two questionnaires will be presented:

The first questionnaire, as already mentioned, asks general questions to put the users into testing groups. Place, date, nationality, age, gender are asked. Moreover, questions about the user's hotel booking and internet experience are asked. This should establish the user's internet and online booking experience and how familiar the user is with both. After extensive brainstorming, this thesis decided on the following questions to understand the user's booking experience (see table 5):

Table 5 – Questions about hotel booking experience

Questions	Possible Answers					
How often do you use the internet?	Never	Monthly	Weekly	Daily	Couple of times a day	Many times a day
How often do you book hotels online?	Never	Once a year	Twice a year	3 - 6 times a year	7 - 12 a times year	>12 times a year
For what purpose do you use computers/notebooks?	Games	Business	E-Shopping	Online Booking	Education	Others
How do you usually book hotels?	Travel Agency	Directly at hotel	Online	I let people book for me	Mobile Application	Others
Which online booking website do you usually use?	Trivago	Expedia	Booking.com	Trip Advisor	HRS	Others

Questions 3 to 5 in the questionnaire above allows more than one answer to be given. There is one essential difference between the first and the second questionnaire. While the first questionnaire asks questions, the second questionnaire gives statements that are based on a Likert-Scale which can be rated between strongly agree (rated by 5) and strongly disagree (rated by 1).

In the second questionnaire, the user gives his opinion and rates his satisfaction with the website. It also asks specific questions to discover weaknesses of the website which may be

concluded in the end. The first six questions are based on John Brooke's (1996) SUS questionnaire which gives an overall user satisfaction of the website. The following 14 questions ask for specific areas that could be improved or are well developed, based on Bastien and Scapin's (1993) criteria. The following table 6 shows all 14 questions and to which criterion each question belongs. The whole questionnaire with all 20 questions can be found in appendix C.

Table 6 – Questions based on Bastien & Scapin (1993) criteria

Nr.	Question	Criteria	Comments
7	I liked the functions offered by the website such as filters.	1. Guidance 1.1 Prompting	Lead the user to make specific actions (e.g. specific entrance or other tasks)
8	The website guided me nicely through the process.	1. Guidance 1.1 Prompting	Lead the user to make specific actions (e.g. specific entrance or other tasks)
9	I always had control in the process. I never felt helpless or clueless.	3. Explicit Control 3.2. User Control	User should be over the control of the process. Actions should be anticipated.
10	The website supported me in completing the task. (e.g. calendar popped up by date entrance)	1. Guidance 1.1 Prompting	Lead the user to make specific actions (e.g. specific entrance or other tasks)
11	There was neither too much nor too little information on the screen.	2. Workload 2.2 Information Density	Provide only necessary and immediately usable data.
12	Text was comfortable and easy to read.	1. Guidance 1.2 Grouping/Distinction of Items	Visual organization of information items.
13	I think the website was efficient. (No unnecessary steps or windows needed)	2. Workload 2.1 Minimal Action	Limiting as much as possible the steps users must go through.
14	I like the labels, symbols, text size and colors on the website.	1. Guidance 1.2 Grouping/Distinction of Items	Visual organization of information items.
15	It was easy to identify the right button I was looking for.	1. Guidance 1.2 Grouping/Distinction of Items	Visual organization of information items.
16	The website provided a good support when problems occurred.	5. Error Management 5.2. Quality of Error Message	Specific error messages, neutral wording, do not blame user.
17	The error messages were very specific.	5. Error Management 5.2. Quality of Error Message	Specific error messages, neutral wording, do not blame user.
18	The error message helped me to correct the error.	5. Error Management 5.3. Error correction	Allow user to edit and re-enter only that portion that must be corrected.
19	The effort required was proportionate to the goal.	2. Workload 2.1. Minimal Action	Limiting as much as possible the steps users must go through.
20	The achieved result met my prior expectations.	3. Explicit Control 3.2 User Control	User should be over the control of the process. Actions should be anticipated.

6.3.2 Quantitative Data Collection

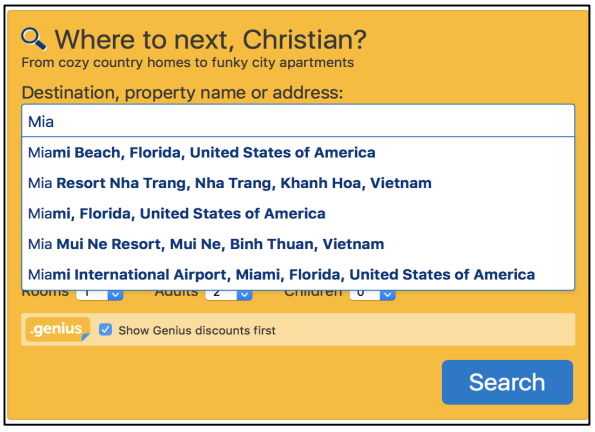
The quantitative or numerical data collection happens during the two tasks executions. To record and document the data collection correctly, the software MORAE© by TechSmith was used as a supporting software. The MORAE© software contains two separate parts, MORAE© Recorder and MORAE© Manager. MORAE© Recorder records all information needed and MORAE© Manager makes it possible to read and analyze this information, as well as re-watching the recorded test.

To collect data correctly, MORAE© Recorder must be set up in the beginning. The two tasks that were defined in chapter 6.1 must be defined to collect the data separately for each task. Moreover, criteria to be recorded must be defined as well. In the following the criteria that are essential for each task will be defined. These criteria are based on Bastien and Scapin (1993). The better these criteria are realized, the more efficient and user friendly the website.

Success or Failure: According to Bastien and Scapin's 6th criteria Error Protection, a good system should prevent the user from failing. Success or failure is probably the most essential criteria and can be answered by yes or no. In the hotel booking process a task has been successfully completed as soon as the user cancelled the hotel reservation. The customer service contacting has been successfully completed as soon as the user has sent the given text to Booking.com.

Keyboard hits to completion: "The shorter the entries, the smaller the probability of making errors" (BASTIEN and SCAPIN, 1993). This is the theory of the Concision criteria. This considers the number of all keys used by the tester. This should focus more on how efficiently a user can type and if he is forced to type unnecessary text. The following case should demonstrate a positive example. Miami Beach is chosen as a destination. The user only must type in three letters and the website automatically gives a suggestion of Miami beach (see figure 12).

Figure 12 – Example of suggestions without typing in complete city (BOOKING.COM, 2017)



Mouse-Clicks to completion: The Minimum Action criteria aims to minimize the number of actions necessary to accomplish the goal or task (BASTIEN and SCAPIN, 1993). Mouse-Clicks to Completion counts the clicks needed to accomplish this task.

Mouse Movement: The Mouse Movement criteria is also adopted by Bastien and Scapin's minimum action criteria. The more user-friendlier the website and the better the understanding of the website, the less mouse movement is needed. If the user is going in circles, it is not efficient at all. If the following step is clear, the target is clear and he can move directly to the right spot.

Windows switch to completion: Each window also requires time to load and time to adapt to the new interface. Again, the Minimum Action criteria of Bastien and Scapin (1993) can be adopted at this point. Consequently, the rule is: The fewer windows needed, the more user-friendlier the website is.

Time to completion: Time to completion is basically the consequence of the last three criteria. Fewer keys used, and fewer clicks and windows needed, may lead to the result in less time. Again, it is based on the Minimal Action criteria but it also comes under Error Protection because the fewer buttons typed or clicks used, the smaller the possibility of making an error which increases the time to completion.

Numbers of Short-cuts: This criterion must be distinguished from Bastien and Scapin's (1993) criteria User Experience and the Minimum-Action-criterion. A short-cut in this evaluation method is any type of action that shortens the entry. On the one hand, a short cut depends on the user's skills, memories, experience etc. Commands like copy-paste depend on the user himself. However, on the other hand, the website can make the usage as simple as possible by providing possible short cuts which leads to the Minimal-Action-criterion. This means, if a destination suggestion appears (see example above on figure 12), it is a short cut provided by the website. The more short-cuts provided by the website but also the more short-cuts possible, the more efficient it makes the website.

Anyway, not every criterion is essential for task 1 and task 2. In the following, criteria for each task are chosen which seem to be significant for the analyses for each task. According to the chosen criteria, MORAE© Recorder must be set up and modified to deliver these numbers for each task. For task 1 the following criteria are defined as essential:

- Success and failure: This is an essential criterion to measure the success rate of the hotel booking process.

- Time: If a problem occurs or the user is inexperienced, he obviously needs more time to complete the task. A long booking process does not necessarily mean a problematic performance. A user can take longer time to book because he is checking reviews, location, pictures etc. of every single option. So, this must be considered before drawing any conclusions.
- Keyboard hits: If a user uses a lot more keyboard hits, the website may be missing supporting features such as suggesting locations when hitting one letter in the destination box. Again, there may be different reasons why a user has hit the keyboard so many times, which must be analyzed as well. Some users may prefer to use the keyboard to switch from one spot to the other and others may prefer to use the mouse click. This is the reason why keyboard and mouse clicks are deeply connected to each other.
- Mouse Clicks: If a user must rebook a hotel or made any mistake, he may need more mouse clicks to complete the task. However, as before, it does not necessarily mean the user did something wrong when he needed many clicks.
- Mouse Movement: The higher the mouse movement, the higher the chance that the user was inexperienced and got lost at some point. He might have been searching for something which was not visible enough. The reason will be analyzed and as mentioned before, the total number will be treated carefully. If a user took more time, he might have moved the mouse more than users that completed the task in a shorter time. To consider this, the mouse movement will be also analyzed by dividing mouse movement for task 1 by its time.

$$\frac{Mouse\ Movement\ task\ 1}{Time\ task\ 1}\ \frac{pixel}{seconds}$$

This relative reference will be also done for Keyboard hits and Mouse Clicks.

- Windows Switched: As soon as a problem occurs, the user must go backwards or open other windows. However, people may analyze the hotel, open reviews etc. which obviously will not count as critical action.
- Short-cut: Short-cuts are very valuable to make usage comfortable and quicker. Here, which user used significantly more short cuts and if the website even made it possible to use short-cuts will be analyzed.

Since the scope of task 2 is different to task 1 and has different purposes, not every criterion that was needed to analyze task 1 is necessary. The main criteria used for task 2 are the following:

- Success or failure is probably the most significant criterion for this task because this shows if the tester was able to contact the customer service or not. The task was successfully completed as soon as the tester sent the text to the customer service. The user could quit the task whenever he wanted to. An exit-criterion was defined to avoid infinity circularizing which ended the task after 7 minutes. By the way, it is unrealistic that a user spends more than 7 minutes searching for the customer service in real conditions.
- Time: Assuming the longer the user needed to find the customer service box, the harder it was to find it.
- Windows Switched: Windows switched is important as well because the MOM to solve this task was exactly one window switch.

6.4 EVALUATION PROCESS

In the previous chapter the data collection was explained. In this chapter, this data will explain how this data will be used to draw conclusions.

As already mentioned, MORAE© was used as a supporting software. The quantitative data collection is completely realized by MORAE© Recorder. The evaluation was realized by MORAE© Manager. This software of the MORAE© package made it possible to compare users to each other, to see differences in each criterion defined above. Unfortunately, not every criterion defined in chapter 6.3.2 can be provided by MORAE©. Time, mouse clicks, mouse movement and keyboard hits can be recorded and exported automatically by MORAE© but success and failure, windows switched and short-cuts must be done manually. MORAE© Manager allows the user re-watch the recorded process in different speed levels which made it possible to extract this data as well.

MORAE© Manager provided a couple of graphs and diagrams, which will be presented in this research as well, but once it comes to specific filters and analysis, MORAE© Manager has some restrictions. To go deeper into analysis, Microsoft© Excel was used as another software that

supports the evaluation. Filters, tables, diagrams and pivot-tables can be easily created by this software. After each test the recorded data was exported to the Excel file. Once all tests were completed, the Excel sheet contained all information collected (as quantitative data, as well as qualitative data). The questionnaires will be filled out by the user directly in the Excel data which allowed Excel to transfer these data automatically to the Excel sheet with all other data collected.

Once all data was collected together in the Excel sheet, the evaluation process could start. First, a quantitative analysis was done which led to defining key users. Key users are users that performed either well or critically. Best performers, as well as worst performers (in this thesis these users are called critical users) are defined for each criterion by each task. This means, for example, the time criterion for task 1 had one best performer who took the least time to complete the booking process and it had one critical user that took the most time. After defining the best users and critical users for each criterion, the evaluation switches to the qualitative evaluation. Reasons for both kinds of performances are searched for. It will be explained why the best performer took less time and why the critical performer took more time.

All problems that occurred with the critical users are summarized and analyzed. How many critical users suffered (and which type of problem they suffered) will be analyzed. Those problems that occurred many times, will be discussed and shown in an Ishikawa Diagram (Cause and Effect Diagram). Originally, the Ishikawa Diagram was created for lean manufacturing and aims the elimination of manufacturing problems. This is also called the 5M model which tries to discover problems in machine, method, material, man and measurement. (LINDEMANN, 2009)

However, this method will be modified for this research and it will not target the 5Ms but rather the weak points identified on the website. Further, the 5-Why method will be applied to the most critical problems. The 5-Why method asks maximum 5 times the question “why does the problem occur?” and each question relates to the previous question. The goal of this method is to discover the reason of the problem. This method was developed by Taiichi Ohno who empirically decided 5 times as a sufficient amount to ask for the reason of the problem. (BICHENO, 2006)

The first why will be answered with at least one criterion of Bastien and Scapin. The lack of those criteria should support the discovery of reason of the problem. All further why-questions will go deeper in detail.

In addition to that, the auto evaluation of each critical user will be considered, as well. The best and the worst rating of each critical user will be noted. The questions with the most negative rating (Frequently Answered Questions), are considered as critical. Since most questions are based on Bastien and Scapin criteria, conclusions and recommendations from the worst rated questions can be drawn.

This evaluation procedure will be undertaken for both tasks. To evaluate the UX factors and difference between user groups during the booking process, other criteria will be defined and analyzed. Each testing group (European vs. Latin America; Brazilian vs. exchange student; age group; male vs. female) will be analyzed by the following criteria:

- Extreme Points: Extreme points are extreme performances. This means it considers the key users. This part will analyze how many users of each testing group were part of the best performance and how many took part of the critical performance.
- Experience: Experience will show, according to the questionnaire at the beginning of the test, which testing groups book hotels online more often and which are more familiar with those booking platforms.
- Time: For each testing group the maximum, minimum and average time needed will be defined.
- Mouse Movement: As with the time criterion, the maximum, minimum and average mouse movement by testing group will be defined.
- Windows Switched: A statistic on how many windows opened to complete task 1 will be shown for each testing group.
- Since Keyboard Hits are strongly connected to Mouse Clicks, they will be considered together. Some users may prefer using the keyboard to scroll up and down or hitting the next button and others may prefer using the mouse to go through the process. To draw conclusions on which testing groups may prefer which option, a graph will be created which shows the relation and division between the testing groups.

6.5 EXECUTION PROCESS

The execution process, realized with every user equally, can be summarized in 10 steps:

1. User fills out a form with basic information on personal details such as age or gender.
2. User fills out a questionnaire about their user experience information regarding online booking experience.
3. User reads both tasks.
4. User starts with task 1 and MORAE© Recorder will be activated (user is asked if he wants to be recorded or not. If he does not want to be recorded, the front webcam will be turned off).
5. After completing task 1, user starts with task 2 (contacting customer service).
6. After finishing task 2, user fills out the 20-question questionnaire (auto evaluation).
7. Testing leader uploads MORAE© Recorder data in MORAE© Manager.
8. Information taken from MORAE© Manager (Time, mouse clicks, mouse movement, keyboard hits, windows switched, short-cuts) and fed into the software based usability evaluation tool.
9. Data is automatically transformed to raw data tab in Excel tool.
10. Next candidate is tested and process starts from the beginning.

6.6 SUMMARY

This chapter has given a deep insight into this usability evaluation method. In chapter 6.1 the task to be carried out by the testers has been defined including different constraints on the hotel booking process. This research formulated hypotheses which will be justified or not in the end. Then, the data collection has been explained in chapter 6.3. Quantitative and qualitative methods have been presented in chapter. In the last chapter, the general execution process, that was repeated with every single tester, has been explained step by step. In the following chapter a software based tool will be presented that supports the evaluation method.

7 SOFTWARE BASED TOOL

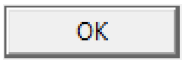
A software based tool was developed to guide the user step by step through the experiment. In the following chapter, the interfaces of this method will be shown.

Since Microsoft© Excel illustrates formulas and diagrams clearly, this software is used to develop this tool. The evaluation and results are illustrated perfectly after executing the whole procedure. Moreover, the VBA program language, offered by Microsoft© Excel, allows for the easy creation of user forms which guide the person, in a user-friendly manner, through the process. The user only must click on specific buttons to express their performance and opinion, as well as feelings, to generate an overall evaluation including diagrams.

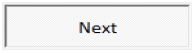
In the first sub-chapter, standard keys that are used in this tool are explained. After this, the step by step procedure, incorporated by user forms, will be presented in chapter 7.2. And since this is a usability evaluation tool, great focus was put on consistency and transparency of this tool. In consideration of this, the last sub-chapter deals with the tool's error prevention.

7.1 STANDARD BUTTONS

Before explaining the evaluation procedure, all standard buttons will be presented in this chapter:



The OK-Button accepts a given statement, shown by the software.



The Next-Button forwards the user to the next step.

A Text Box allows the user to type in a free text. Some text boxes have restrictions, such as numbers only.

Toggle-Buttons allow the user to select more than one answer option.

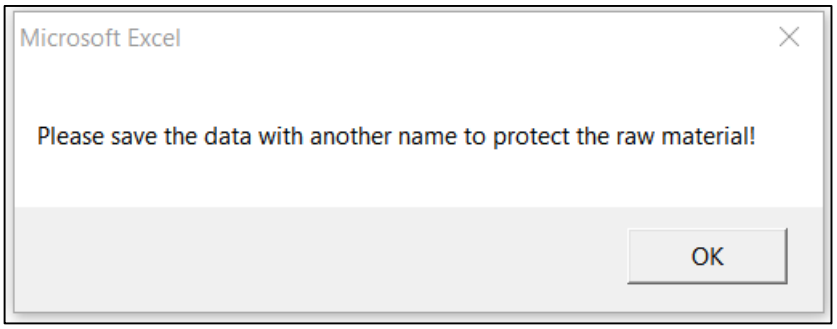
At least two option buttons appear per question/statement. Only one option button can be activated per question/statement.

A Drop-Down shows possible answers. One of the offered answers must be chosen. Drop-Downs are useful if many options are possible such as country selection.

7.2 STEP BY STEP PROCEDURE

Before even opening the software based tool, the tool automatically asks the user to change the name of the data which automatically creates a copy of the data. This protects the raw data from any unwanted changes (see figure 13). To activate this feature, the original file must be saved under the name “Evaluation Method”.

Figure 13 – Message to protect raw data



After saving the file under a different name and copying the file successfully, the first tab will be opened. On this tab the testing procedure can be started by clicking on the start-button or a completely new test can be started by clicking on the delete-button which deletes all data collected before (see figure 14).

If the user decides to start a new process the first user form appears that asks for personal information such as age, gender and nationality (see figure 15).

And the second user form will ask questions about the user’s internet and hotel online booking experience (see figure 16). The option buttons of the first two question on the internet and hotel online booking experience form only allow the tester/user to choose one of the given answers. The last three questions allow the user to answer more than one question by using toggle-buttons.

Figure 14 – Starting interface

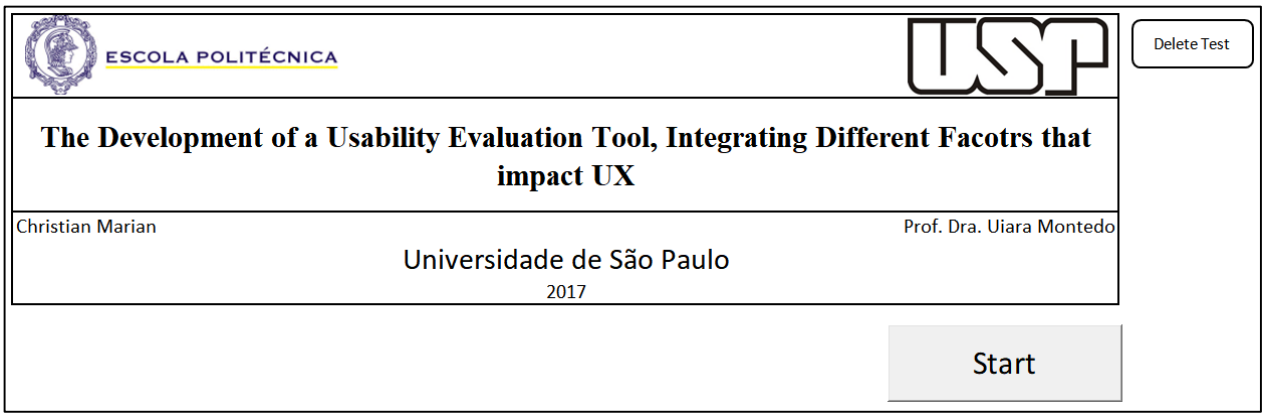


Figure 15 – Personal information form

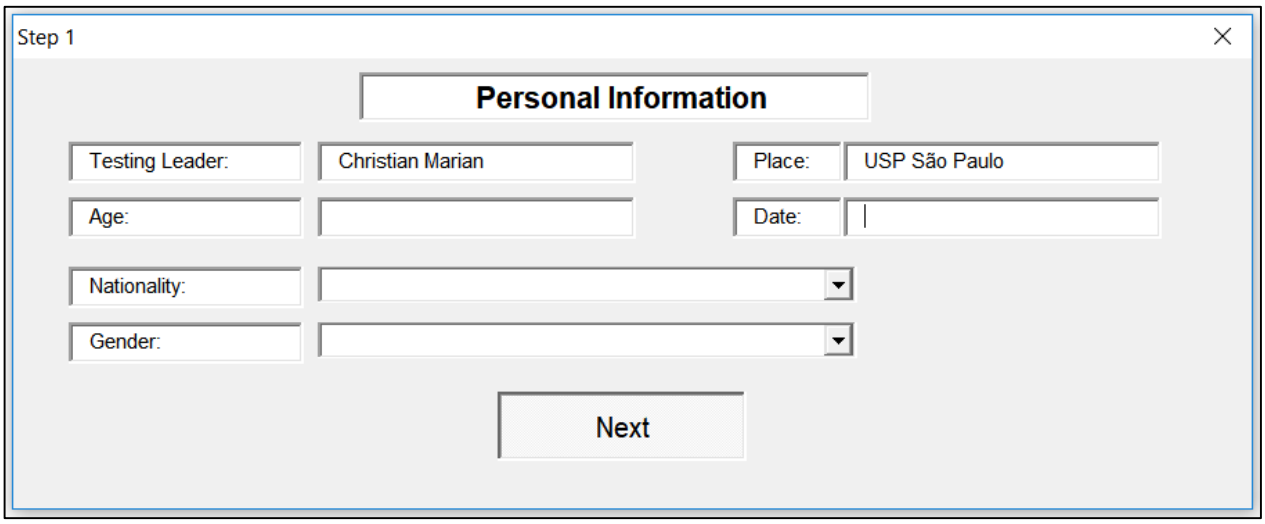
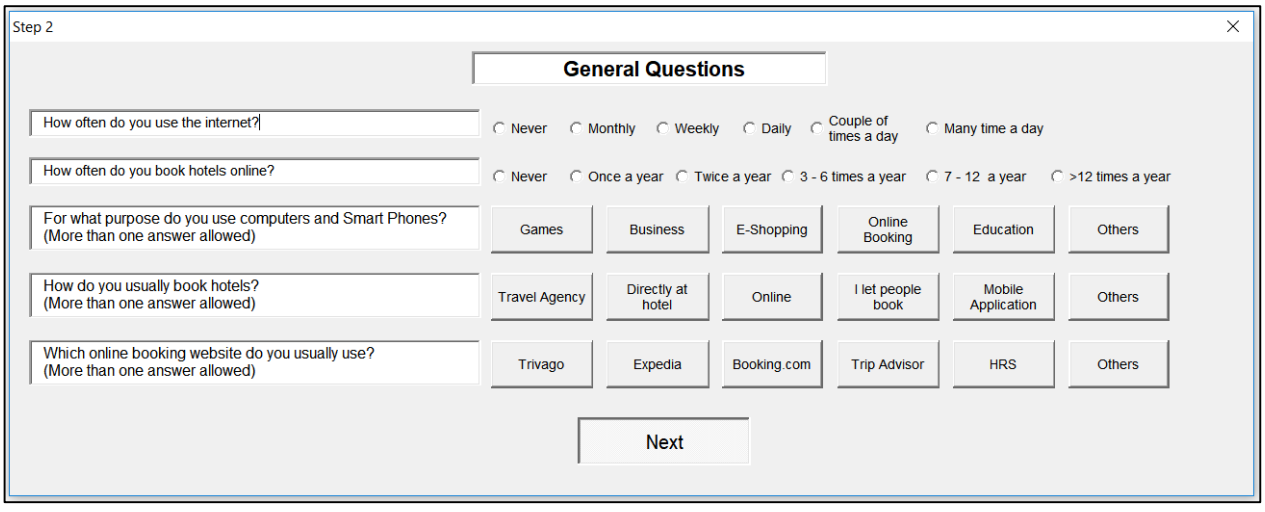


Figure 16 – Internet and online booking experience form



After collecting all general information, the user is forwarded to a user form which explains the hotel booking task (tasks that were defined in chapter 6.1 on page 64 and 65). However, the user does not have to remember all constraints. A printed version of this task will be given to the user to look up the constraints at any time during the test. The text that must be sent to the customer support in the end, will be shown in a separate window to give the opportunity to use a copy-paste function. After accepting and understanding the task, the user will be automatically forwarded to the booking website.

After completing the task, the user will close the internet window and a questionnaire will appear immediately. This questionnaire was already presented in chapter 6.3.1 and will be presented here digitally in figure 17.

After filling out the questionnaire, the tester's job is completed. A user form will appear which expresses appreciation for his effort and asks him to let the testing leader fill out the following form shown in figure 18. This form will ask for the numerical performance. Since this developing tool records the user performance such as time, clicks, mouse movement etc., this form will collect this data from MORAE© Manager which is manually fed into this Excel user form.

Figure 17 – Questionnaire form

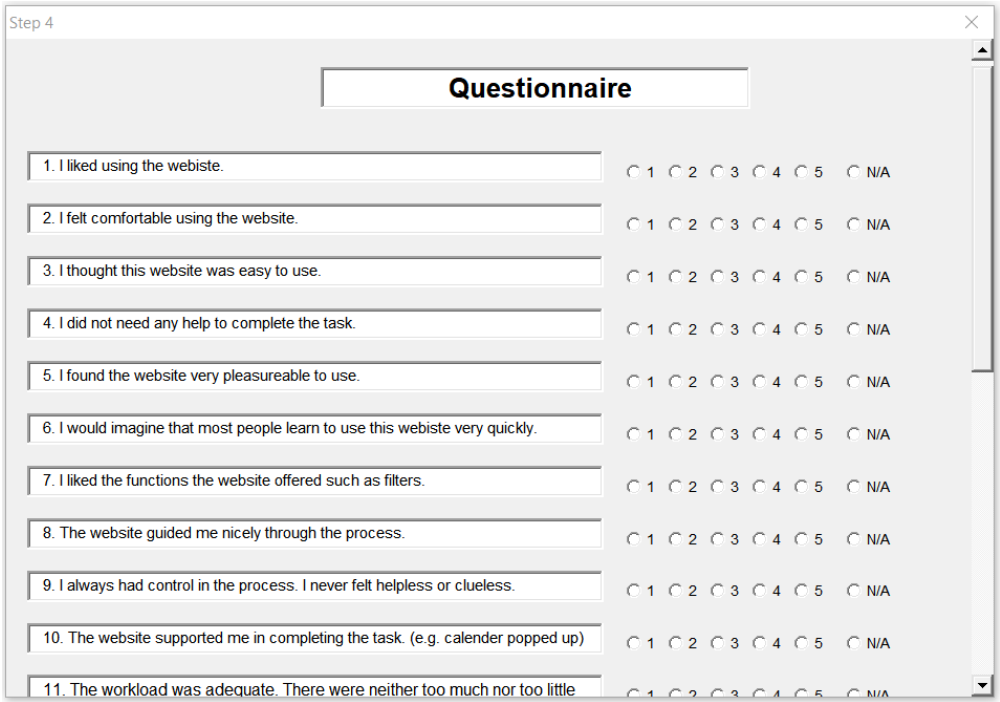
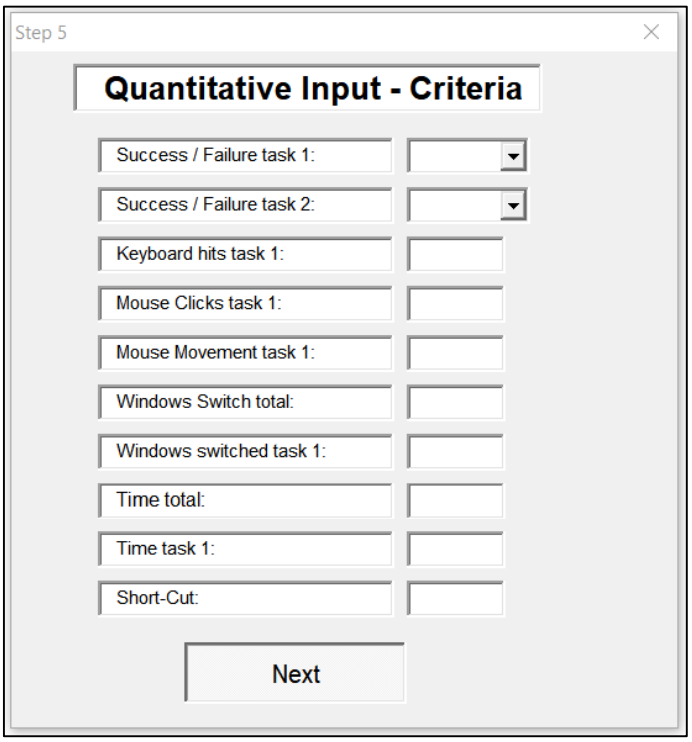


Figure 18 – Numerical performance documentation



Notice that only the total time and time for task 1 was documented because time for task 2 can be easily calculated with those two number. The same counts for the windows switched. This choice was taken to avoid errors and to save time.

$$Total\ Time\ [seconds] - Time\ Task\ 1\ [seconds] = Time\ Task\ 2[seconds]$$

After this form, all data is collected. From this point, the collected data can be evaluated and interpreted. All answered questions and all input can be seen on the Raw Data tab. Excel forwards the testing leader automatically to this tab.

To repeat the test with another person, either the Start-button on the raw data tab or on the starting tab must be activated.

The raw data tab is not supposed to be changed. It is only an overview of the data collection. The number of testers can be seen easily because every new tester will be added to a new line in this worksheet. Furthermore, filters can be set to do quick evaluations. The actual analyses are done on further tabs. Diagrams and pivot-tables show the desired information concerning the testing groups.

7.3 ERROR PREVENTION

Since user-friendliness and transparency are important to this software based tool, a separate chapter about error prevention was created. Some examples of error prevention are given, such as age consistency or numerical input check.

Since the testing environment has been defined as a student environment, the software asks the user to check their input as soon as he types in an unrealistic age. For example, if the user puts in a three-digit age, which is not possible, an error message appears and the user cannot continue with the process until he has corrected the error. Figure 19 shows this error exemplary message.

Figure 19 – Age inconsistency error message

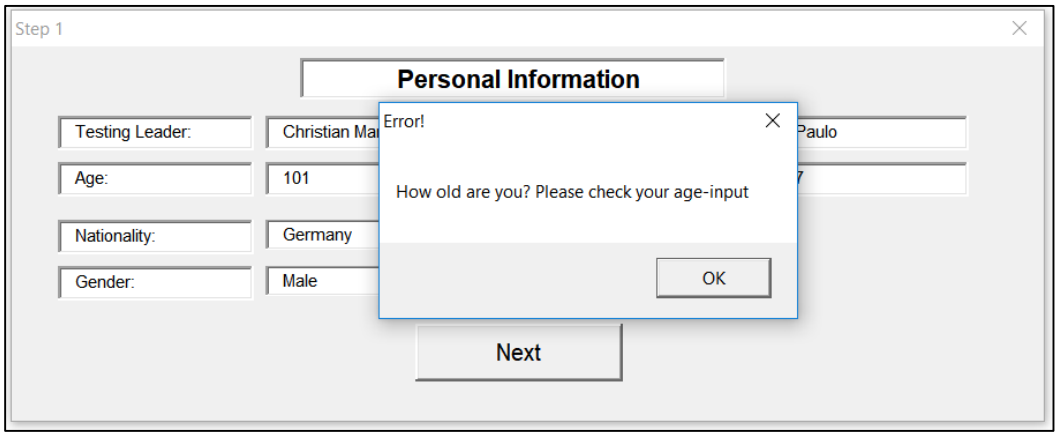


Figure 20 gives an example of an error message if an obligatory box is empty. Figure 21 shows what happens if an incorrect input was fed in. Since this text box must contain numerical input, the following error message would appear if it is not the case.

7.4 SUMMARY

This chapter has presented a software-based evaluation tool. This was developed by Microsoft© Excel. In the beginning, all important buttons were illustrated and explained. In the second step, all user forms were presented including the raw data and analysis tabs in the end. In the last sub-chapter, some examples of error preventions error messages were given.

Figure 20 – Error message due to empty text box

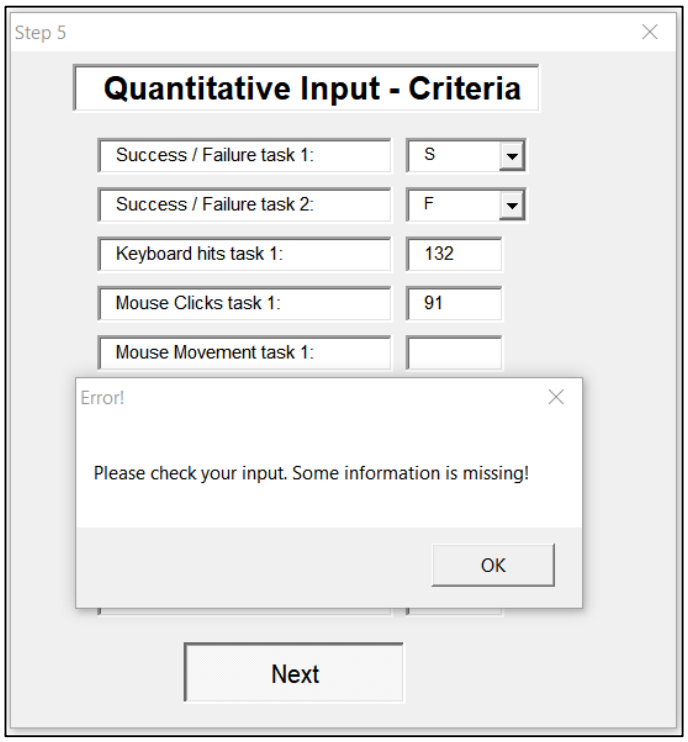
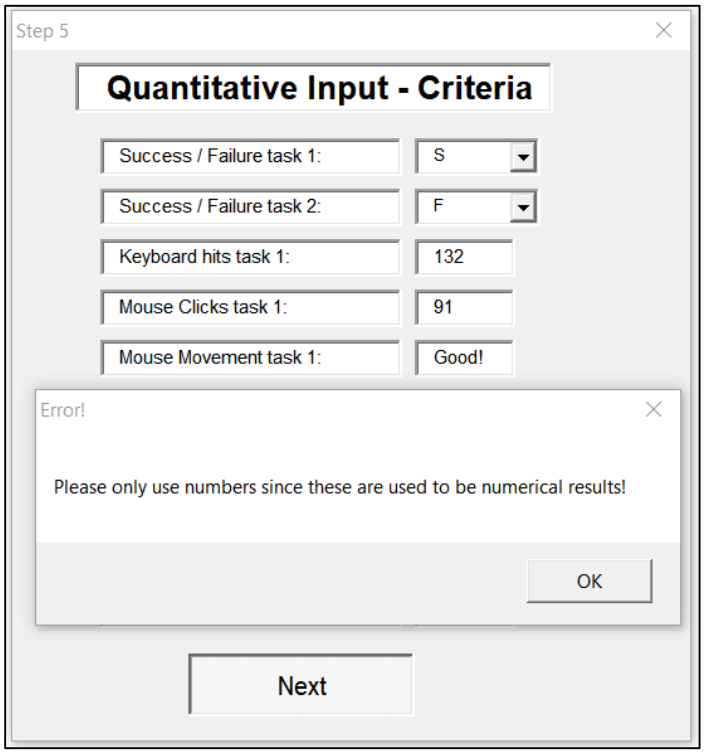


Figure 21 – Error message due to non-numerical input



8 APPLICATION OF USABILITY TOOL

The application of the presented usability evaluation tool was executed on September 11th till September 14th 2017 in the OCEAN Samsung© lab in the *Escola Politecnica* of the University of São Paulo. The used equipment was a Samsung Notebook NP600B4C.

In the following chapter the results will be presented. First, an overview of the output of MORAE© and Excel will be given. In the second part (chapter 8.2), an overview of the participating persons will be given. An exemplary division will be shown. Finally, the results and qualitative numbers that can be drawn from the test by each testing group will be given.

8.1 RESULTS AND ANALYZES BY SOFTWARE

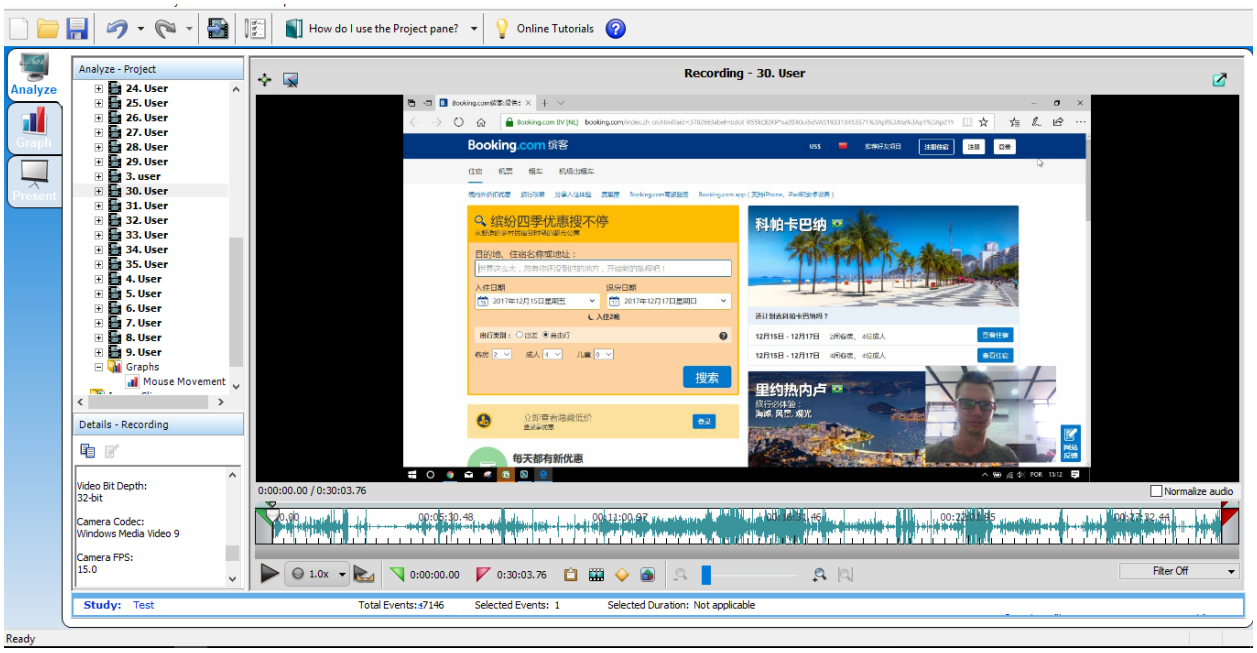
As explained in chapter 6.3.2, the supporting software that were used were MORAE© by TechSmith and Excel by Microsoft©. Basically, MORAE© was used to record and deliver the quantitative information and Excel was used to analyze that information.

Figures 22 – 28 will illustrate how all necessary criteria and data (time, mouse clicks, keyboard hits, mouse movement, windows switched and short-cuts) were collected by MORAE© Manager. Figure 22 shows, randomly, how user 30 started the booking process (Mandarin Booking.com home page). By clicking the start button the whole booking procedure can be watched again at different speed levels. This was an important and powerful feature because reasons for problems that occurred could be analyzed and observed after the testing.

Time, mouse clicks, keyboard and mouse movement could be exported by MORAE© Manager directly to Excel. The data was exported firstly, numerically as a list and secondly, as a graph. To get a better overview, the graphs that were created by MORAE© Manager in Excel will be shown in the following figure 23 – 26.

As already explained in chapter 6.4 on page 72, mouse clicks, keyboard hits and mouse movement were only considered for task 1. The following graphs will illustrate these data only for this task.

Figure 22 – Exemplary MORAE Manager starting interface



The following figure will show how the time data (total time, time 1 and time 2) was collected by MORAE.

Figure 23 – Time in seconds for entire recording, task 1 and task 2

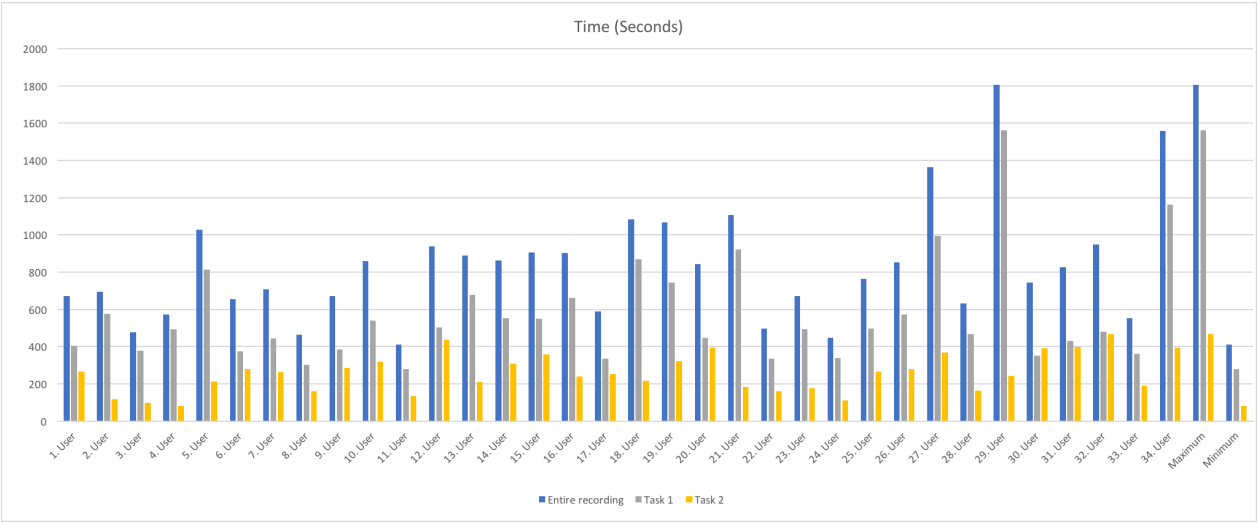
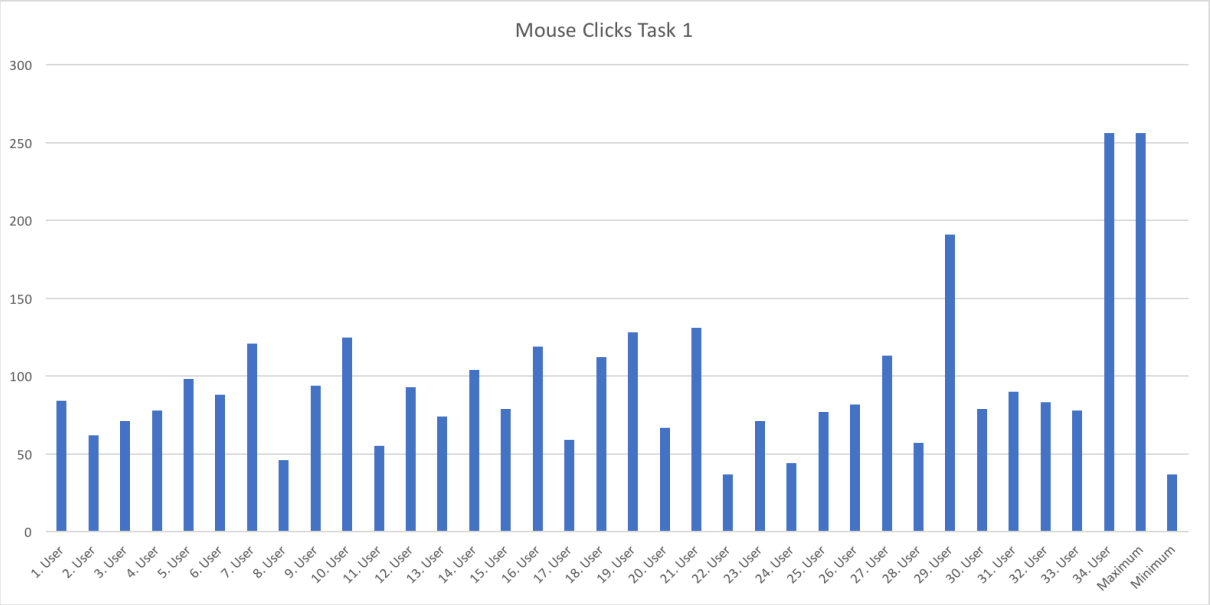


Figure 24 – Mouse clicks task 1



The graph above (figure 24) shows the mouse click for task one per user and the figure below shows the keyboard hits per user for task 1.

Figure 25 – Keyboard hits task 1

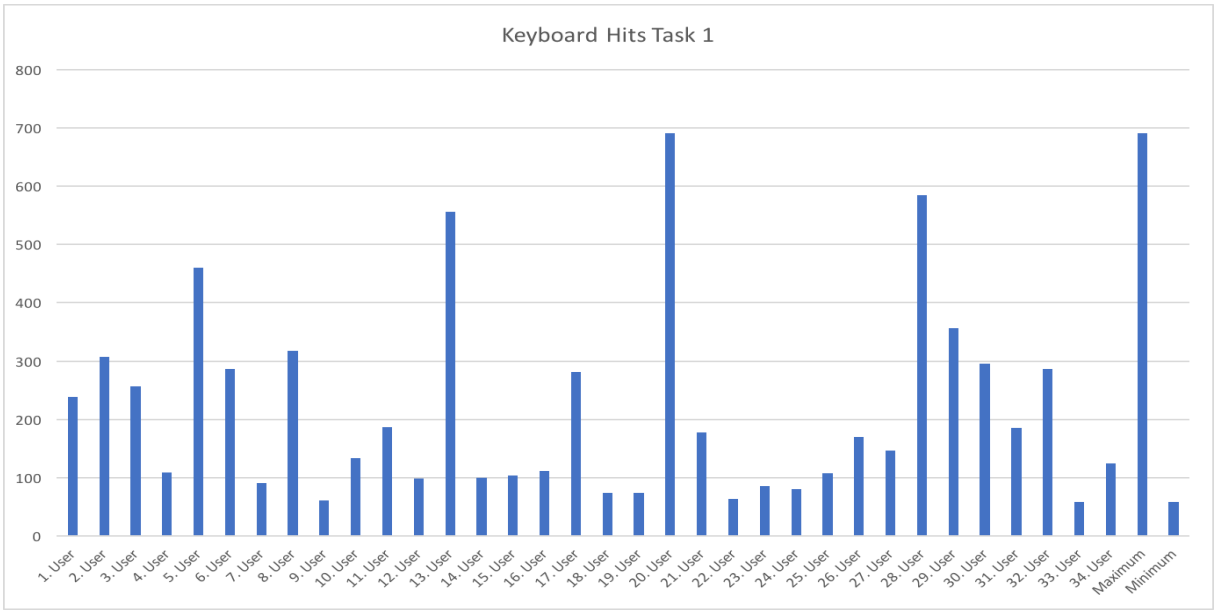
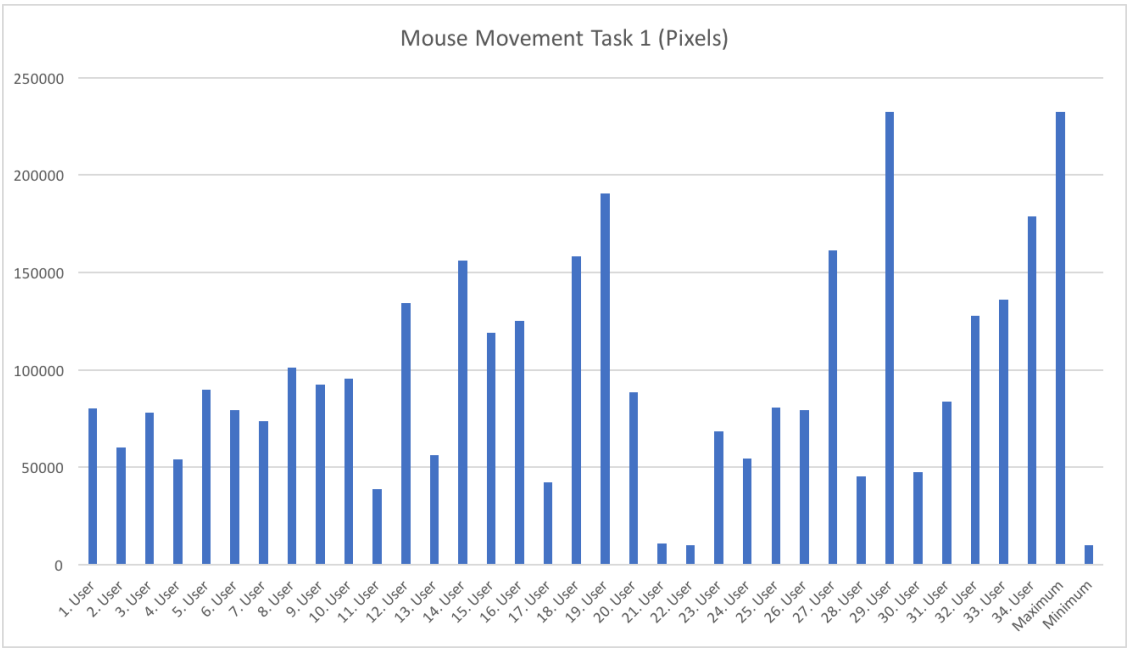


Figure 26 – Mouse movement in pixel for task 1



Since there was no way to count windows and short-cuts automatically by MORAE©, this had to be done manually. Additionally, these two graphs that were created manually by Excel will be shown in figure 27 and 28.

Figure 27 – Short-cuts entire recording

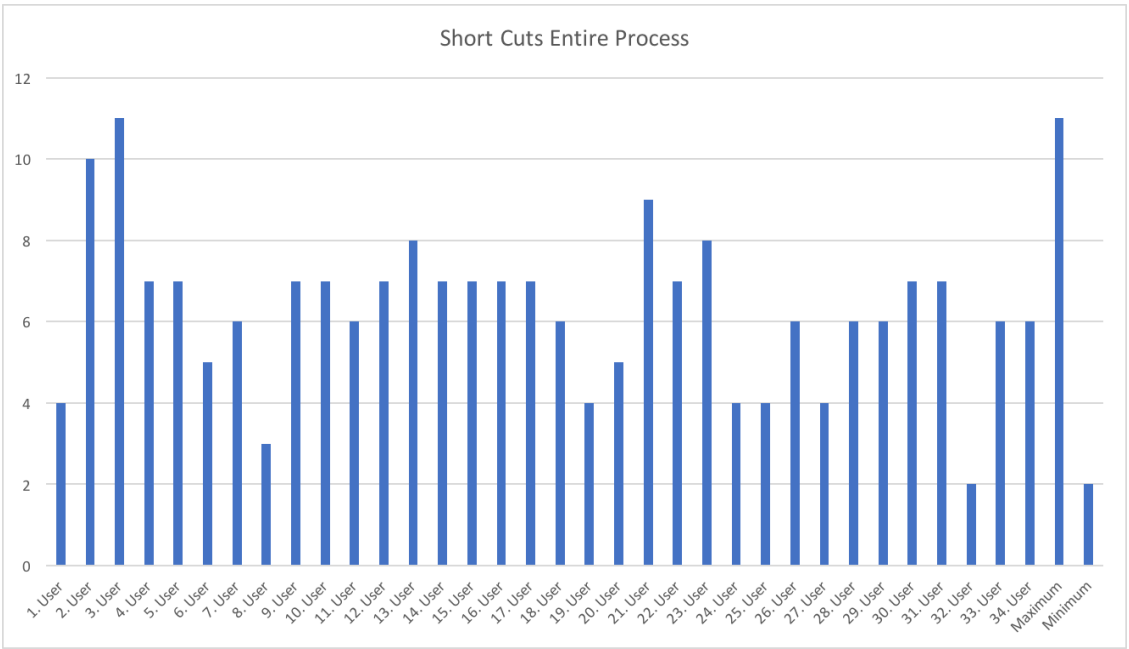
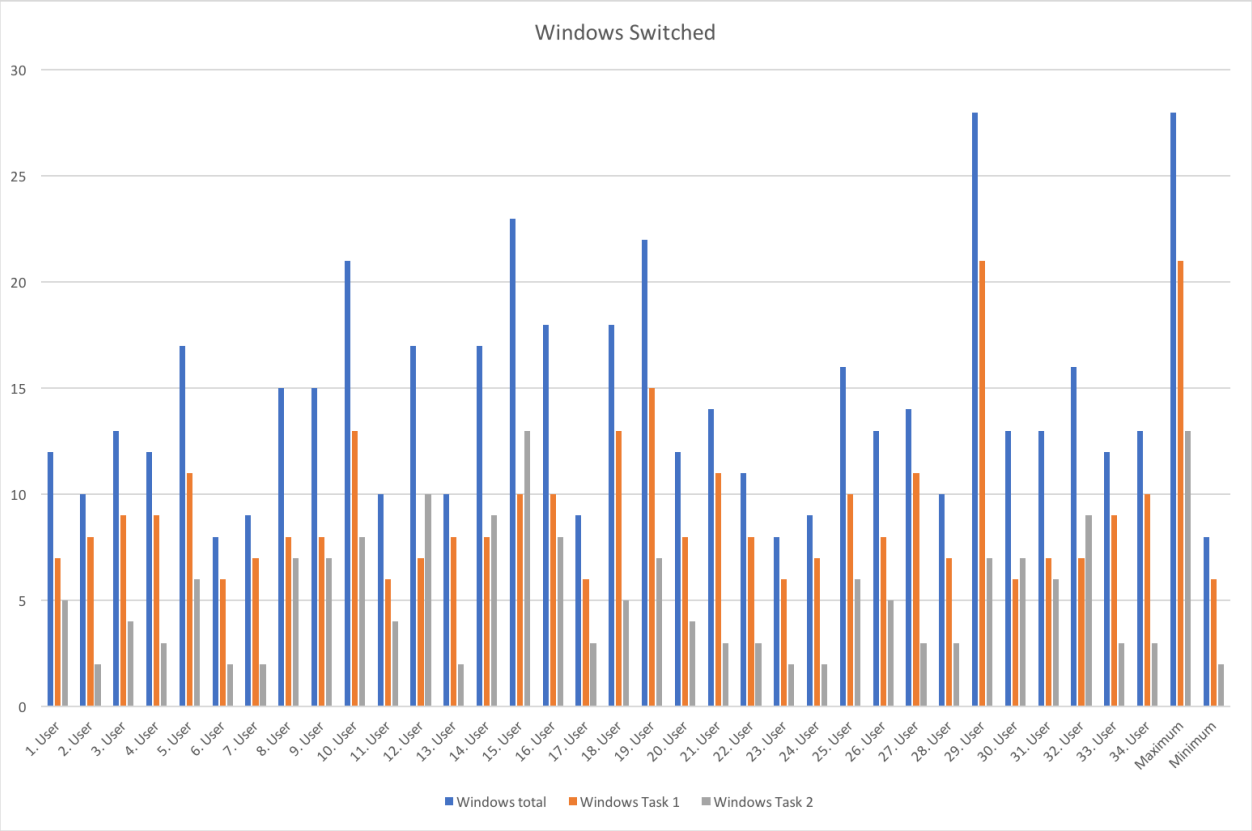


Figure 28 – Windows switched entire task, task 1 and task 2



As the MORAE© software and its output has been explained, the Microsoft© Excel file will be explained step by step. As already presented in chapter 7.2 on page 81, after every tester, the Excel tool's form allowed the forms to be filled out confirming the data collected by MORAE©. After filling out this form, it was automatically fed into the raw data sheet. Furthermore, all information provided by the user (general information like age, nationality, gender etc.), as well as the answers of two questionnaires, were fed in into this sheet. Figures 29 to 31 show the complete raw data sheets with all information entered.

In the following, the whole raw data sheet is explained. Basically, this sheet consists of six different parts:

1. User Information: This section gives general information about the user (User Number, date of testing, place of testing, age, gender and nationality). It is basically filled in the form shown in figure 15 in chapter 7.2 on page 79.

2. Auto Evaluation: This part of the raw data shows the questionnaire filled in by each user after the test, keep in mind that question 1 to 6 were questions about the general satisfaction and question 7 to 20 are based on Bastien and Scapin's criteria. With this in mind, averages have been established, to give a quick overview of the user's overall and criteria-based satisfaction.
3. Collected data: This is basically the data collected by MORAE. This data was filled in in the form that was presented in figure 18 in chapter 7.2 on page 81.
4. Experience Questionnaire: Here all answers given by the user on the questionnaire, that had been handed in before the booking process started, can be found (see figure 16 in chapter 7.2 on page 79). Note, a 1 in the last 3 columns (Purpose, How book hotel? And Booking.com) means that the user considers the internet as an online booking platform (purpose column), books hotels online when he needs to book one (How book hotel?) and uses Booking.com when he books a hotel (Booking.com).
5. Relative Data: Relative data are basically the data keyboard hits, mouse clicks and mouse movement for task 1 divided by time for task 1.
6. Testing Groups: These columns divide the users into different testing groups (defined by UX factors and hypotheses in chapter 6.2 on page 66) which helps to filter the data in the following analyzes. A "+" in the key user column symbolizes a best performer, a "-" symbolizes a critical performer and a 0 means no key user.

The first 4 parts were always filled out automatically by the Excel tool's forms. Parts 5 and 6 were automatically created by Excel formulas.

Figure 29 – Raw data sheet part 1

User Information						Auto Evaluation																				
User	Date	Place	Age	Gender	Nation	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	Av	Av. 1-6	Av 7-20
1	11.09.2017	USP São Paulo	24	Male	Italy	5	4	5	4	5	5	5	5	4	5	3	4	5	3	3	5	5	4	4	5	4,40	4,67	4,29
2	11.09.2017	USP São Paulo	17	Male	Brazil	3	3	3	4	4	3	3	3	2	4	4	4	3	3	3	4	4	4	4	2	3,35	3,33	3,36
3	11.09.2017	USP São Paulo	22	Male	Brazil	2	2	2	4	2	2	3	3	2	4	2	3	2	2	2	4	4	4	3	4	2,80	2,33	3,00
4	12.09.2017	USP São Paulo	23	Male	Brazil	2	2	2	4	3	2	5	3	2	5	1	1	1	3	1	3	4	N/A	3	4	2,68	2,50	2,77
5	12.09.2017	USP São Paulo	18	Male	Brazil	4	4	5	3	4	5	5	5	4	5	5	5	3	4	3	N/A	N/A	N/A	4	5	4,29	4,17	4,36
6	13.09.2017	USP São Paulo	23	Male	Brazil	3	2	2	1	3	2	5	3	2	4	3	5	4	5	2	N/A	N/A	N/A	4	5	3,24	2,17	3,82
7	13.09.2017	USP São Paulo	24	Male	Brazil	5	5	5	3	5	5	4	4	4	4	4	4	4	4	4	4	N/A	N/A	4	4	4,22	4,67	4,00
8	13.09.2017	USP São Paulo	24	Female	Brazil	3	2	3	4	3	2	4	2	3	2	3	4	4	3	3	N/A	N/A	N/A	4	3	3,06	2,83	3,18
9	13.09.2017	USP São Paulo	18	Male	Brazil	5	5	5	4	5	4	5	4	4	5	4	5	5	5	5	5	N/A	N/A	5	5	4,72	4,67	4,75
10	13.09.2017	USP São Paulo	24	Female	Italy	4	4	4	4	3	3	5	4	4	4	4	4	5	4	4	4	4	4	4	4	4,00	3,67	4,14
11	13.09.2017	USP São Paulo	23	Male	France	4	4	4	5	4	3	4	4	4	4	5	4	4	4	5	3	3	3	2	3	3,80	4,00	3,71
12	13.09.2017	USP São Paulo	23	Male	Finland	4	4	3	3	4	4	5	4	2	3	2	5	3	5	2	1	3	2	3	3	3,25	3,67	3,07
13	13.09.2017	USP São Paulo	21	Female	Peru	4	5	5	4	5	5	5	4	5	3	2	3	4	4	4	4	5	5	4	4	4,20	4,67	4,00
14	13.09.2017	USP São Paulo	23	Male	Colombia	2	2	4	1	2	2	2	2	2	2	2	3	3	2	3	2	4	3	2	2	2,35	2,17	2,43
15	13.09.2017	USP São Paulo	22	Female	Italy	4	3	4	2	3	4	5	3	2	3	2	4	4	3	3	3	4	3	4	4	3,35	3,33	3,36
16	13.09.2017	USP São Paulo	18	Male	Brazil	4	3	3	1	3	2	4	3	2	4	2	3	4	4	2	3	1	2	4	2	2,80	2,67	2,86
17	13.09.2017	USP São Paulo	21	Male	Germany	5	5	5	5	5	5	5	5	4	5	3	4	5	5	4	5	5	5	5	4	4,70	5,00	4,57
18	13.09.2017	USP São Paulo	21	Female	Venezuela, RB	2	2	1	1	1	5	5	1	1	2	2	3	1	3	1	2	2	1	1	1	1,90	2,00	1,86
19	13.09.2017	USP São Paulo	19	Female	Colombia	4	4	3	4	3	4	5	4	4	4	5	4	5	3	3	4	N/A	N/A	5	5	4,06	3,67	4,25
20	13.09.2017	USP São Paulo	22	Female	Colombia	4	4	3	3	4	4	5	3	3	2	3	4	4	4	4	3	4	4	3	4	3,60	3,67	3,57
21	14.09.2017	USP São Paulo	25	Male	Brazil	3	3	2	5	2	3	5	1	2	3	3	3	2	3	2	N/A	N/A	N/A	4	5	3,00	3,00	3,00
22	14.09.2017	USP São Paulo	24	Male	Italy	4	4	5	5	4	4	5	4	5	4	4	4	5	4	4	N/A	N/A	N/A	4	5	4,35	4,33	4,36
23	14.09.2017	USP São Paulo	26	Male	Germany	4	4	4	5	4	3	5	4	4	5	3	3	4	4	4	N/A	N/A	N/A	5	5	4,12	4,00	4,18
24	14.09.2017	USP São Paulo	22	Male	Germany	4	4	3	2	3	5	4	4	4	2	2	5	4	3	4	N/A	N/A	N/A	5	5	3,71	3,50	3,82
25	14.09.2017	USP São Paulo	21	Male	Brazil	4	4	3	2	4	3	5	4	4	4	2	3	3	4	4	3	N/A	N/A	4	4	3,56	3,33	3,67
26	14/09/2017	USP São Paulo	21	Male	France	4	4	5	3	4	5	4	5	5	4	4	5	5	5	3	4	4	5	4	5	4,35	4,17	4,43
27	14/09/2017	USP São Paulo	25	Female	Germany	3	4	4	2	4	4	5	5	3	4	4	5	2	2	4	4	4	5	5	5	3,90	3,50	4,07
28	14/09/2017	USP São Paulo	21	Male	Mexico	5	5	5	5	4	4	5	4	4	5	5	5	5	3	4	5	5	5	5	5	4,65	4,67	4,64
29	14/09/2017	USP São Paulo	23	Male	Germany	3	3	4	5	3	4	4	3	3	4	2	4	2	3	3	2	N/A	N/A	4	4	3,33	3,67	3,17
30	14/09/2017	USP São Paulo	23	Male	Mexico	4	4	5	2	4	5	5	4	2	5	3	4	4	5	3	4	4	3	5	3	3,90	4,00	3,86
31	14/09/2017	USP São Paulo	21	Female	Colombia	4	3	4	3	4	4	5	5	4	4	5	5	4	4	4	3	5	N/A	3	4	4,05	3,67	4,23
32	14/09/2017	USP São Paulo	20	Female	Mexico	4	5	5	3	5	3	5	5	4	5	5	5	5	5	4	5	5	5	5	5	4,65	4,17	4,86
33	14/09/2017	USP São Paulo	22	Female	Germany	4	5	4	5	4	5	5	4	5	5	5	5	5	4	4	3	5	5	5	5	4,60	4,50	4,64
34	14/09/2017	USP São Paulo	32	Female	Brazil	2	2	3	1	2	1	4	3	2	4	3	4	2	3	2	1	1	1	2	1	2,20	1,83	2,36

Figure 30 – Raw data sheet part 2

Collected Data											
Success / Failure Task 1	Success / Failure Task 2	Keyboard	Mouse Clicks	Mouse Movement	Windows	Windows Task 1	Windows Task 2	Time	Time Task 1	Time Task 2	ShortCuts Task 1
S	F	238	84	80342,69	12	7	5	671,17	405	266,17	4
S	F	307	62	60044,01	10	8	2	693,9	576	117,9	6
S	F	256	71	78220,94	13	9	4	478,61	379	99,61	7
S	S	109	78	54109,33	12	9	3	573,02	492	81,02	7
S	S	460	98	89943,23	17	11	6	1027,71	814	213,71	7
S	S	287	88	79199,44	8	6	2	653,88	374	279,88	5
S	S	91	121	73717,16	9	7	2	708,02	444	264,02	6
S	F	317	46	101068,15	15	8	7	464,55	304	160,55	3
S	F	61	94	92270,16	15	8	7	671,64	386	285,64	7
S	F	134	125	95381,88	21	13	8	858,63	540	318,63	7
S	F	186	55	10091,98	10	6	4	412,06	278	134,06	6
S	S	98	93	134241,64	17	7	10	938,05	502	436,05	7
S	F	556	74	56250,68	10	8	2	888,69	679	209,69	8
F	S	100	104	156135,07	17	8	9	863,65	553	310,65	7
S	F	104	79	118878,72	23	10	13	906,04	549	357,04	7
S	F	112	119	125246,24	18	10	8	901,65	661	240,65	7
S	F	281	59	42397,53	9	6	3	589,19	337	252,19	7
S	F	74	112	158382,82	18	13	5	1084,61	868	216,61	6
S	F	74	128	190451,13	22	15	7	1067,15	745	322,15	4
S	S	691	67	88524,20	12	8	4	844,18	448	396,18	5
S	F	177	131	11092,02	14	11	3	1107,28	923	184,28	9
S	F	63	37	39079,98	11	8	3	496,27	337	159,27	7
S	F	85	71	68376,67	8	6	2	671,1	493	178,1	10
S	F	80	44	54705,01	9	7	2	448,58	338	110,58	4
S	F	107	77	80659,87	16	10	6	765,1	498	267,1	4
S	F	170	82	79579,83	13	8	5	852,69	572	280,69	6
S	F	146	113	161220,19	14	11	3	1363,06	996	367,06	4
S	F	584	57	45538,44	10	7	3	630,61	468	162,61	6
S	F	357	191	232419,51	28	21	7	1803,69	1560	243,69	6
S	S	296	79	47643,62	13	6	7	745,59	353	392,59	7
S	S	185	90	83777,11	13	7	6	828,07	431	397,07	7
S	F	286	83	127867,26	16	7	9	947,19	481	466,19	2
S	F	58	78	136067,44	12	9	3	551,67	362	189,67	6
S	S	125	256	178998,45	13	10	3	1559,54	1164	395,54	6

Figure 31 – Raw data sheet part 3

Experience Questionnaire					Relative Data *(divided by time task 1)*			Testing Groups				
Internet Frequency	Online Booking Frequency	Purpose	How book hotel?	Booking.com	Keyboard	Mouse Clicks	Mouse Movement	Continent	Brazil / Exchange	BRA/EUR/LA	Age Range	Key User
Many times	7-12 times a year	1	1	1	0,59	0,21	198,38	Europe	Exchange	Europe	>22 years old	0
Many times	Never	0	0	0	0,53	0,11	104,24	Latin America	Brazil	Brazil	<20 years old	0
Many times	Never	1	1	1	0,68	0,19	206,39	Latin America	Brazil	Brazil	20 - 22	0
Many times	2-6 times a year	1	1	1	0,22	0,16	109,98	Latin America	Brazil	Brazil	>22 years old	0
Many times	Once a year	0	0	0	0,57	0,12	110,50	Latin America	Brazil	Brazil	<20 years old	0
Many times	Once a year	0	1	0	0,77	0,24	211,76	Latin America	Brazil	Brazil	>22 years old	+
Many times	7-12 times a year	1	1	1	0,20	0,27	166,03	Latin America	Brazil	Brazil	>22 years old	0
Many times	Once a year	0	1	1	1,04	0,15	332,46	Latin America	Brazil	Brazil	>22 years old	0
Many times	Twice a year	0	1	1	0,16	0,24	239,04	Latin America	Brazil	Brazil	<20 years old	0
Many times	More than 12 times a year	1	1	1	0,25	0,23	176,63	Europe	Exchange	Europe	>22 years old	0
Many times	Twice a year	0	1	1	0,67	0,20	36,30	Europe	Exchange	Europe	>22 years old	+
Many times	2-6 times a year	1	1	1	0,20	0,19	267,41	Europe	Exchange	Europe	>22 years old	0
Many times	Once a year	1	1	1	0,82	0,11	82,84	Latin America	Exchange	Latin America	20 - 22	0
Many times	Never	0	1	0	0,18	0,19	282,34	Latin America	Exchange	Latin America	>22 years old	-
Many times	7-12 times a year	1	1	1	0,19	0,14	216,54	Europe	Exchange	Europe	20 - 22	0
Many times	Once a year	0	1	1	0,17	0,18	189,48	Latin America	Brazil	Brazil	<20 years old	-
Many times	Once a year	1	1	1	0,83	0,18	125,81	Europe	Exchange	Europe	20 - 22	+
Many times	Never	1	0	0	0,09	0,13	182,47	Latin America	Exchange	Latin America	20 - 22	-
Many times	2-6 times a year	0	1	1	0,10	0,17	255,64	Latin America	Exchange	Latin America	<20 years old	-
Many times	2-6 times a year	0	1	0	1,54	0,15	197,60	Latin America	Exchange	Latin America	20 - 22	0
Many times	Once a year	0	0	1	0,19	0,14	12,02	Latin America	Brazil	Brazil	>22 years old	0
Many times	2-6 times a year	0	1	1	0,19	0,11	115,96	Europe	Exchange	Europe	>22 years old	+
Many times	7-12 times a year	1	1	1	0,17	0,14	138,70	Europe	Exchange	Europe	>22 years old	+
Many times	Twice a year	1	1	1	0,24	0,13	161,85	Europe	Exchange	Europe	20 - 22	0
Many times	2-6 times a year	0	1	0	0,21	0,15	161,97	Latin America	Brazil	Brazil	20 - 22	0
Many times	Once a year	0	1	1	0,30	0,14	139,13	Europe	Exchange	Europe	20 - 22	0
Many times	More than 12 times a year	1	1	1	0,15	0,11	161,87	Europe	Exchange	Europe	>22 years old	0
Many times	Twice a year	0	0	0	1,25	0,12	97,30	Latin America	Exchange	Latin America	20 - 22	0
Many times	7-12 times a year	1	1	1	0,23	0,12	148,99	Europe	Exchange	Europe	>22 years old	-
Many times	Once a year	0	1	1	0,84	0,22	134,97	Latin America	Exchange	Latin America	>22 years old	+
Many times	Once a year	0	1	0	0,43	0,21	194,38	Latin America	Exchange	Latin America	20 - 22	0
Many times	Once a year	0	1	0	0,59	0,17	265,84	Latin America	Exchange	Latin America	20 - 22	0
Many times	Twice a year	1	1	1	0,16	0,22	375,88	Europe	Exchange	Europe	20 - 22	0
Many times	Once a year	0	0	1	0,11	0,22	153,78	Latin America	Brazil	Brazil	>22 years old	-

8.2 PARTICIPANT TESTER

In total 34 persons were tested. 12 Brazilians and 22 exchange students participated in this research. From these exchange students, 9 were Latin Americans and 13 from Europe. Six out of 13 Europeans were German. There were 12 female testers and 22 male testers. 5 out of 34 were teenagers and 16 testers were older than 22 years-old. This age range was chosen regarding the representative number of participants of each age range. In the following table 7, an exemplary division of testing groups is shown. In brackets the number of testers of this specific group is listed.

Table 7 – Exemplary testing group division

<table>
<tr><td rowspan="4">Europe (13)</td><td>Germany (6)</td><td rowspan="2">Male (9)</td><td>< 20 years old (0)</td></tr>
<tr><td>Italy (4)</td><td rowspan="2">20 – 22 years old (5)</td></tr>
<tr><td>France (2)</td><td rowspan="2">Female (4)</td></tr>
<tr><td>Finland (1)</td><td>>22 years old (8)</td></tr>
<tr><td rowspan="5">Latin America (21)</td><td>Brazil (12)</td><td rowspan="2">Male (13)</td><td rowspan="2">< 20 years old (5)</td></tr>
<tr><td>Mexico (3)</td></tr>
<tr><td>Colombia (4)</td><td rowspan="3">Female (8)</td><td>20 – 22 years old (8)</td></tr>
<tr><td>Peru (1)</td><td rowspan="2">>22 years old (8)</td></tr>
<tr><td>Venezuela (1)</td></tr>
</table>

8.3 RESULTS

In this chapter results in numbers will be discussed. In the first part, the results of task 1 will be revealed. The second part targets the second task and in the end, the results of the three chiosen UX factors are presented.

8.3.1 Results Task 1

In this sub-chapter, the results of the hotel booking process, which is designed to evaluate the core function of the website, will be presented.

In chapter 6.3.2 (page 69), criteria have been defined that evaluate task 1 quantitatively (time, keyboard hits, mouse clicks, mouse movement, windows switched and short-cuts).

In the following, each criterion, starting with time, will be presented. Key users – critical and best performer – will be defined regarding their performance numbers. In the meantime, reasons for their outstanding numbers will be examined by re-watching the process on the MORAE© Manager software. A deep analysis and conclusions of these reasons will be discussed in following chapter 9.

The user that took the longest time was user 29 with 1560 seconds to complete task 1. Analyzing his performance, he did fail at one point but this failure was not the only reason for the long duration. He missed a filter which forced him to do the whole procedure again because a credit card was asked at the final booking procedure step. He spent the most time on analyzing different hotel options. He also switched to different websites to read reviews about those options. However, this user did fail and took the longest, which makes him a critical user.

The next critical user is user 14. This was the only user, that failed on the booking process. Frustrations made this user to give up after 553 seconds. The main reason for this were problems with meeting the requirements. He did not have any control over the process through the entire time.

The quickest user was user 11 with 278 seconds for task 1. He was very self-confident, did not have problems at any point of the booking process and found the cancellation button right away. He is the first defined best performer in this research.

To get more critical users and to get more data about issues on this website, critical user will be defined concerning the booking experience (best performer in those categories will not be rewarded because these may manipulate further statistics, while critical performers were defined to find weak points of the website and improve it). Since six different options have been given to describe someone's online booking frequency (never, once a year, twice a year, between 2 and 6 times a year, 7 to 12 times a year or more than 12 times a year), every single option is analyzed by the number of times.

In total 4 users never book hotels online. Critical user 18 took 868 seconds to complete the first task. This user had problems with filters. It took a while to figure out how to filter the constraints correctly. He was insecure and seemed to be lost. His dissatisfaction is also shown by his auto evaluation which gave an overall average of 1,9 (out of 5!).

Among the users that book online once a year. User 34, who took 1160 seconds for task 1, seemed to struggle and will be part of the critical user group, as well. User 34 was overwhelmed using filters. He did not have control over the process at any time of the booking and was relieved after completing the task. This user also admitted in the end, that he usually books journeys directly at travel agencies.

Considering users that book hotels between 2 and 6 times a year, the critical user that took the longest time was user 19. This user took 745 seconds. He had problems in the beginning with changing the language but also with using the filters. It took a while till he found the right filters. First he scrolled up and down and tried to find the right hotels chaotically.

Among users that book a hotel more than 6 times a year, no significant problem occurred, except user 29 who missed one filter in the beginning and had to rebook. Anyway, this user has already been grouped with the critical users in the beginning of this chapter.

Mouse Movement

Mouse movement will be the second criteria which is analyzed. First, the user that used the mouse most economically was user 22 with 10,091.98 pixels. User 29 is considered as the user with the highest mouse movement (232,419.51 pixels). As already mentioned in the paragraph above, a problem indeed occurred, but another reason for moving the mouse so much was his booking procedure in general. He switched to other websites, opened different tabs with different options and analyzed them. The second highest mouse movement was from user 19 with 19,0451.13 pixels. User 19 was already discussed in time criteria. He was sometimes lost and scrolled up and down to find a hotel manually first, without using the right filters.

Another interesting analytical comparison is mouse movement in relation to time. Users that took more time, might spend more time moving the mouse. To put this into relation, dividing mouse movement by time can bring interesting observations. Doing this, two exceptions will

appear. User 8 and user 33 lead this measurement by far with 217 pixel/s and 246 pixel/s . The reason is not any significant problem or issue that occurred during the process. The reason is that they moved the mouse with their eyes. The other users looked at the spot they want to click on and then they moved to the spot. Since no problems occurred they will not be critical users but this observation will be considered later in this research.

Windows Switched

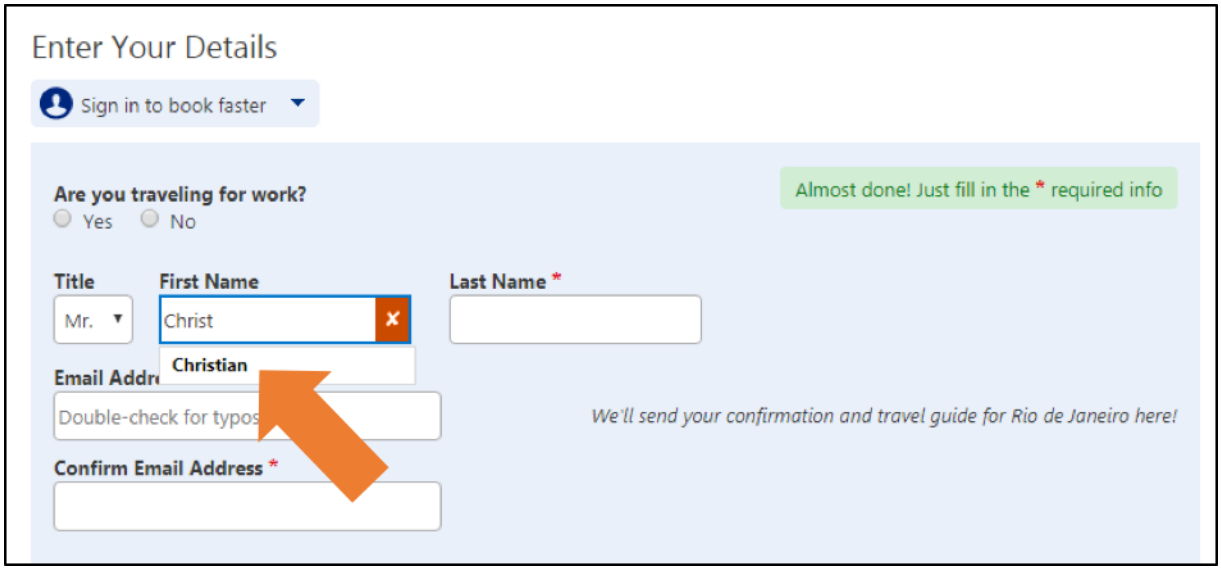
Many users, that did not have any problems with the booking procedures, needed either 6 or 7 windows to complete the task (bearing in mind that 6 windows belong to the MOM of windows switched). Five users only needed 6 windows and seven users needed 7 windows to switch to complete it. The five most efficient performers, who will be rewarded as best performers in this category and will belong to the key users, were users 6, 11, 17, 23 and 30.

One critical user was user 29 with 21 switched windows. However, he was already discussed in detail in the time criteria. User 19 switched around 15 times to complete the task. Reasons for using almost three times more windows than the best performer, were problems finding the language button. He missed couple of windows at this point but also because he tried out different hotels without filtering and booking systematically. The third highest windows switched performance goes to user 10 who switched in total 13 times. However, this user booked the hotel successfully and confidently without any issues. The reason for the high number of windows switched was a deep analysis of different options. He opened different tabs with different hotels and looked at them in detail. Since he neither performed as the best in this category nor experienced any problems he will not be considered as a key user. User 18, in contrast, did have problems. He also switched windows 13 times but he failed the first time, so he had to go back to the hotel list and start the procedure from the beginning. The reason for failing the first time was missing a filter.

Short-Cuts

The standard number of short-cuts varied between 6 and 7. In total 20 users used either 6 or 7 short-cuts during the process. The most common short-cuts were the copy paste function or using suggestions given by the website, such as Copacabana in the destination text box or names remembered from the booking history in the personal information window. (see figure 32).

Figure 32 – Frequent short-cut example



The user with the fewest short-cuts was user 2 with 2 short-cuts. This user completed the task without any problems even though he seemed to be a little inexperienced. He was the only user who looked at the keyboard while typing. Consequently, he missed most suggestions given by the website. The person with most short-cuts (user 23 with 10 short cuts) completed the task successfully, as well. He seemed to be very confident, switched windows with short-cuts and even searched for the breakfast filter with the short cut that led to the search function.

Even though it may be interesting to look at short-cuts and which users used which short-cuts, it does not help to find the weak points of the website. Consequently, it will not be considered in the further research.

Mouse Clicks and Keyboard hits

Due to the strong connection between mouse clicks and keyboard hits (see chapter 6.3.2 page 72), these two criteria will be considered together. To get useful key users on this criterion, users were chosen that used keyboard or/and mouse clicks significantly over or under the average. The average number of keyboard hits was 213 hits per user and the average of mouse clicks were 92 clicks. User 22 was the best performer in this category because in both criteria, keyboard and mouse clicks, he is far under the average (63 keyboard and 37 mouse clicks). This user completed the task straight forwardly and very efficiently, as already mentioned above. User 34 received attention even though he used fewer keyboard hits than the average. However, he

used 256 clicks which is significantly over the average. As already mentioned before, this user had difficulties with the booking process. Besides that, this user was supposed to scroll down with the mouse. Instead he clicked on the scrolling bar each time he wanted to scroll up or down which is the reason for many mouse clicks. User 29 also clicked the mouse more than average with 191 times and the keyboard more than 357 times. As already figured out, this user failed the first time but since he also analyzed the hotels on different websites, he had to type in lots of letters in the new tabs. User 16 also received attention with 119 clicks, even though he only hit the keyboard 112 times. However, one reason for using almost 20% more clicks than others, might be his insecurity. He did not find the filters the first time and had to rebook because the first hotel did not conform to the constraints. This makes him to another critical user.

A summarized overview of each criteria with all key users can be found in the table 8.

8.3.2 Results Task 2

Significant criteria for the second task, as already discussed in chapter 6.3.2 on page 73, are success or failure, time and windows switched.

24 testers were not able to complete task 2. They either gave up or they spent more than 7 minutes on the task. 10 people found the customer service box which represent 29.41% of all testers. Regarding the time, by far the fastest user who found the customer service box did it in 81 seconds, which was the only user under 200 seconds. He only switched the window 3 times and found it relatively quickly. The user who tried the hardest, but still did not find it, was user 32. He reached the 7-minute exit-criterion and was asked to quit after reaching this time. He had no clue, was desperate in the end even though he was only a few clicks away from the customer service box. User 3 already gave up after 99.61 seconds because of getting angry and depressed. He clicked around and after not finding the wanted box at tabs like “contact us” or “feedback”, he quit.

Regarding the windows switched, one user even reached 13 different windows. He was chaotic and clicked around and hoped to reach the goal. He was close to his goal, being only one click away from opening the customer service box, like many other users, as well.

Table 8 – Overview of key user

Criteria	Type of Key User	User	Comment
Time	Critical Performer	User 29	Missed filter and had to rebook. Analyzed options and read reviews.
		User 14	Insecure, lost, no control and did not know how to use filters.
		User 18	High dissatisfaction, seemed to be lost, did not know how to use filters.
		User 34	Insecure, did not know how to user filter.
		User 19	Problems with finding the language button and using the filters.
	Bester Performer	User 11	Confident, no problems and straight forward.
Mouse Movement	Critical Performer	User 29	He had to rebook but also analyzed options, switched to other websites to read reviews.
		User 19	Scrolled up and down, looked for hotels without structure because no filter usage in the beginning.
	Bester Performer	User 22	Efficient and confident.
Windows Switched	Critical Performer	User 29	Due to failure of first run and switching to other websites to analyze hotel.
		User 19	Due to problems of finding language button and failure of first run because did not know how to filter in beginning.
		User 18	Due to failure of first run and rebooking.
	Bester Performer	User 6	Changed only 6 windows to complete task 1. This equals to the MOM of this task.
		User 11	
		User 17	
		User 23	
		User 30	
Keyboard / Mouse Clicks	Critical Performer	User 35	Scrolled up and down with mouse.
		User 29	Typed in other website URL to get in on other website to analyze hotels.
		User 16	20% mouse clicks because of rebooking.
	Bester Performer	User 22	Efficient usage and straight forward.

8.3.3 Results regarding UX Factors

In this chapter results between selected UX factors were revealed. At the beginning of the development of this method the UX factors were limited to culture, age and gender. The following table 9 illustrates this division once again with the number of participating persons in each category in brackets.

Based on the developed evaluation process in chapter 6.4 (page 73), the results of the six criteria (extreme points, experience, time, windows switched, mouse movement and keyboard vs. mouse clicks) regarding the three testing group divisions will be presented. Remembering that culture was divided once into continents and into exchange students vs. Brazilian students. In the following, for each criterion a table will be presented which will be followed by an explanation.

Table 9 – Testing groups division

Culture		
Latin America *(21)*	Europe *(13)*	
Brazilian *(12)*	Exchange Students *(22)*	
Gender		
Male *(22)*	Female *(12)*	
Age Group		
<20 years old (5)	22 – 22 years old (13)	>22 years old *(16)*

Extreme Points

To analyze the extreme points, chapter 8.3.1 will be taken as a reference and key users for the online booking process will be considered.

Once the statistics were given, absolute and relative comparison between the groups were done. In the further tables, absolute will be shortened by *abs*. and relative by *rel*.

Extreme Points by Continent

Table 10 – Key user by continent

Key User by Continent

Continent	Critical	Best	Total	% abs. Critical	% abs. Best
Europe	1	4	13	7,69	30,77
Latin America	5	2	21	23,81	9,52
Total	**6**	**6**			
Europe % rel.	16,67	66,67			
Latin America % rel.	83,33	33,33			

Considering these facts on table 10, 83.33% of the critical users came from Latin America and only 16.67% from Europe. In contrast, 66.67% users with best performances came from Europe and only 33.33% came from Latin America.

Indeed, more Latin Americans have been tested, which increases the chance of having more absolute critical and best performances. To have a relative comparison between those two groups, a percentage of how many users of each testing group were critical and how many were best performer. Only one European out of 13 Europeans shows up in the critical performance list. This means 7.69% of Europeans are classed as critical users. In contrast, 5 out of 21 Latin Americans are critical users which means 23.81% of all Latin Americans showed a weak performance. Considering the best performances, 30.77% of Europeans (4 out of 13 Europeans) got this award and less than 10% of Latin Americans impressed with their performance (2 out of 21 Latin American).

Extreme Points by criterion Brazilian vs. Exchange Student

According to table 11, 4 exchange students were part of the critical users (in total 6 critical user) and only 2 Brazilians are grouped as this category. This means 66.67% were exchange students that performed critically and only 33.33% Brazilians performed badly. In other numbers, 18.18% of all exchange students performed critically and 16.67% of all Brazilians performed critically. 22.73% of all exchange students performed well and only 8.33% of all Brazilians were included in the best performances. Relatively, 83.33% of all good performers were exchange students and only 16.67% were Brazilians.

Table 11 – Key user by Brazilian vs. exchange student

Key User by Brazilian vs. Exchange students

Brazil or Exchange	Critical	Best	Total	% abs. Critic	% abs. Best
Exchange	4	5	22	18,18	22,73
Brazilian	2	1	12	16,67	8,33
Total	**6**	**6**			
Exchange % rel.	66,67	83,33			
Brazilian % rel.	33,33	16,67			

These number may be deceiving, because Europeans were mixed with Latin Americans in the exchange group. If the exchange group is split into European exchange students and Latin American exchange students, the numbers look different. Table 12 shows the numbers.

Table 12 – Key user by Brazilian vs. European and Latin American exchange students

Key User by Brazilian vs. European and Latin American exchange students

Brazil or Exchange	Critical	Best	Total	% abs. Critic	% abs. Best
Brazil	2	1	12	16,67	8,33
Europe	1	4	13	7,69	30,77
Latin America	3	1	9	33,33	11,11
Total	**6**	**6**			
Brazil % rel.	33,33	16,67			
Europe % rel.	16,67	66,67			
Latin America % rel.	50,00	16,67			

In this statistic, every third Latin American exchange student performed critically and only 7.69% of Europeans among the critical users. In contrast, almost every third Europeans was part of the best performance group and only 11.11% of Latin Americans were included. Relatively it also makes a difference because almost 2/3 of all best performers came from Europe and every second critical user was a Latin American exchange student.

Since such a big difference is shown between European and Latin American exchange students (short LA), from now on the European and Latin American exchange students will be regarded separately.

Extreme Points by Age Group

Table 13 – Key user by age group

Key User by Age Group

Age	Critical	Best	Total	% abs. Critical	% abs. Best
<20 years old	2	0	5	40,00	0,00
20 - 22	1	1	13	7,69	7,69
>22 years old	3	5	16	18,75	31,25
Total	**6**	**6**			
<20 years old % rel.	33,33	0,00			
20 - 22 years % rel.	16,67	16,67			
>22 years old % rel.	50,00	83,33			

Table 13 shows the extreme performance by age ranges. It will be quickly clear that no teenager is part of the best performance group and 40% of all teenagers were part of the critical users group. Considering the relative numbers only 1/3 of all critical users were teenagers and every second user was older than 22 years. Almost every third person older than 22 years performed well and less than one in five performed critically.

In this case, absolute and relative numbers differ much because more people from the older age group were tested than teenagers.

Extreme Points by Gender

Table 14 – Key user by gender

Key User by Gender

Gender	Critical	Best	Total	% abs. Critical	% abs. Best
Female	3	0	12	25,00	0,00
Male	3	6	22	13,64	27,27
Total	**6**	**6**			
Female % rel	50,00	0,00			
Male % rel.	50,00	100,00			

Considering the gender in table 14, exactly 50% of the critical performers were female and 50% were male. However, there were almost twice as many male testers as female testers. Consequently, 3 out of 12 women (25%) performed critically and only 13.64% of male testers performed critically. In contrast, all users that performed well were male and none was female. In absolute numbers, it means, 27.27% of all men belong to the best performer group.

Experience

Experience may differ between European and Latin American, exchange students and Brazilian students, between different age groups or between male and female users. Questions were defined in chapter 6.3.1 on page 68. to analyze the user's experience and how familiar they are with online booking. The following tables illustrate the results regarding the different testing groups.

Experience by Continent

Table 15 – Experience by continent

Continent

Hotel booking per year

Hotels booked / year	Europe	Latin America	% Europe	% Latin America
More than 12 times	2	0	15,38	0,00
7 to 12 times	4	1	30,77	4,76
2 to 6 times	2	4	15,38	19,05
Twice	3	2	23,08	9,52
Once	2	10	15,38	47,62
Never	0	4	0,00	19,05
Total	**13**	**21**		

Internet purpose

Internet for booking?	Europe	Latin America	% Europe	% Latin America
No	3	16	23,08	76,19
Yes	10	5	76,92	23,81
Total	**13**	**21**		

How do you book hotels?

How booking hotels?	Europe	Latin America	% Europe	% Latin America
Not online	0	6	0,00	28,57
Online	13	15	100,00	71,43
Total	**13**	**21**		

Which websites do you use to book hotels?

Booking.com Yes/No?	Europe	Latin America	% Europe	% Latin America
No	0	10	0,00	47,62
Yes	13	11	100,00	52,38
Total	**13**	**21**		

According to table 15, almost every second European books a hotel online more than 6 times a year and only 4.76% of Latin Americans book a hotel online more than 6 times. However, almost 60% of all Latin Americans book a hotel online maximum once a year (19.05% never book online and 47.62% book once a year). And there was no European who does not book any hotels online a year and only 15.38% book a hotel once a year.

These numbers also reflect on the purpose of the internet. 3/4 of all Europeans said that one purpose of the internet is online booking and the same percentage of Latin Americans answered the opposite. Going deeper into this question, all Europeans, when booking hotels, do so online. In contrast, 71.43% of all Latin Americans do so.

Furthermore, every European uses Booking.com as an online booking platform and only every second Latin American (52.38%) stated that Booking.com is one option for them to book hotels online.

Experience by Brazilian vs. Exchange students

Assuming exchange students travel more and have more necessity to book hotels than a Brazilian who lives in his home town or country, the statistics will be analyzed as follows:

Table 16 – Experience by Brazilian vs. European and Latin American exchange students

Brazilian vs. European and Latin American exchange students

Hotel booking per year

Hotels booked / year	Brazil	Europe	Latin America	% Brazilian	% Europe	% LA
More than 12 times	0	2	0	0,00	15,38	0,00
7 to 12 times	1	4	0	8,33	30,77	0,00
2 to 6 times	2	2	2	16,67	15,38	22,22
Twice	1	3	1	8,33	23,08	11,11
Once	6	2	4	50,00	15,38	44,44
Never	2	0	2	16,67	0,00	22,22
Total	**12**	**13**	**9**			

Internet purpose

Internet for booking?	Brazil	Europe	Latin America	% Brazilian	% Europe	% LA
No	9	3	7	75,00	23,08	77,78
Yes	3	10	2	25,00	76,92	22,22
Total	**12**	**13**	**9**			

How do you book hotels?

How booking hotels?	Brazil	Europe	Latin America	% Brazilian	% Europe	% LA
Not online	4	0	2	33,33	0,00	22,22
Online	8	13	7	66,67	100,00	77,78
Total	**12**	**13**	**9**			

Hotel booking per year

Booking.com Yes/No?	Brazil	Europe	Latin America	% Brazilian	% Europe	% LA
No	4	0	6	33,33	0,00	66,67
Yes	8	13	3	66,67	100,00	33,33
Total	**12**	**13**	**9**			

Brazilians as well as Latin Americans never book hotels more than 12 time a year, while 15.38% of Europeans do. Furthermore, almost half of all European exchange students book more than 6 hotels a year. In contrast, no Latin American does. Only 8.33% of Brazilian students book between 7 and 12 hotels a year. Moreover, 2/3 of all Latin Americans book a hotel maximum once a year. Besides that, only 22.22% of all Latin Americans book their hotels online and only every third person uses Booking.com as one online booking platform.

In contrast, every European books frequently on Booking.com. Also, most Brazilians do not book more than 1 hotel a year (66.67%) but when they book a hotel online, 2/3 of all Brazilians use Booking.com. However, only 1/4 of them consider the internet as an online booking purpose.

Experience by Age Group

Table 17 – Experience by age group

Age Group

Hotel booking per year

Hotels booked / year	<20 years old	20 - 22 years old	>22 years old	% <20 years old	% 20 - 22 years old	% >22 years old
More than 12 times	0	1	1	0,00	7,69	6,25
7 to 12 times	0	1	4	0,00	7,69	25,00
2 to 6 times	1	2	3	20,00	15,38	18,75
Twice	1	2	2	20,00	15,38	12,50
Once	2	5	5	40,00	38,46	31,25
Never	1	2	1	20,00	15,38	6,25
Total	**5**	**13**	**16**			

Internet purpose

Internet for booking?	<20 years old	20 - 22 years old	>22 years old	% <20 years old	% 20 - 22 years old	% >22 years old
No	5	6	8	100,00	46,15	50,00
Yes	0	7	8	0,00	53,85	50,00
Total	**5**	**13**	**16**			

How do you book hotels?

How booking hotels?	<20 years old	20 - 22 years old	>22 years old	% <20 years old	% 20 - 22 years old	% >22 years old
Not online	2	2	2	40,00	15,38	12,50
Online	3	11	14	60,00	84,62	87,50
Total	**5**	**13**	**16**			

Which websites do you use to book hotels?

Booking.com Yes/No?	<20 years old	20 - 22 years old	>22 years old	% <20 years old	% 20 - 22 years old	% >22 years old
No	2	6	2	40,00	46,15	12,50
Yes	3	7	14	60,00	53,85	87,50
Total	**5**	**13**	**16**			

Considering the three age groups, every fifth teenager never books hotels online. And 80% book maximum two hotels online a year. 50% of all tester older than 22 years book at least more than 2 hotels a year. Almost every third person (31.25%) books more than 6 hotels a year. Almost 70% of people between 20 and 22 years old book no more than 2 hotels a year, but 7.69% book more than 12 times a year.

No teenager considers the purpose of the internet to be for booking hotels, although 60% of them answered that, in the case of needing to book a hotel, they would do so online. In contrast, 80% of both older age groups consider that they would use the internet if it was necessary to book a hotel. 60% of all teenagers consider Booking.com as one of their online booking platforms, 53.85% of the testers between 20 and 22 years do so and 87.50% of the oldest age group consider this website as one of their online booking platforms.

Experience by Gender

Table 18 – Experience by gender

Gender

Hotel booking per year

Hotels booked / year	Female	Male	% Female	% Male
More than 12 times	2	0	16,67	0,00
7 to 12 times	1	4	8,33	18,18
2 to 6 times	2	4	16,67	18,18
Twice	1	4	8,33	18,18
Once	5	7	41,67	31,82
Never	1	3	8,33	13,64
Total	**12**	**22**		

Internet purpose

Internet for booking?	Female	Male	% Female	% Male
No	6	13	50,00	59,09
Yes	6	9	50,00	40,91
Total	**12**	**22**		

How do you book hotels?

How booking hotels?	Female	Male	% Female	% Male
Not online	2	4	16,67	18,18
Online	10	18	83,33	81,82
Total	**12**	**22**		

Which websites do you use to book hotels?

Booking.com Yes/No?	Female	Male	% Female	% Male
No	4	6	33,33	27,27
Yes	8	16	66,67	72,73
Total	**12**	**22**		

Considering table 18, 50% of all female testers do not book a hotel online more than once a year. The male testers displayed similar numbers with 45.45%. Two out of 12 women answered that they book more than 12 hotels a year. No male tester answered this. However, 8.33% of female testers book a hotel online 7 – 12 times a year and 18.18% of all males do too. 50% of all women consider a purpose of the internet to be online booking and 40,91% of all men agree with this. 83.33% of all female testers book their hotels online and 81.82% of males do so. Finally, 66.67% of all women use Booking.com as an online booking platform and 72.73% of all men.

Time

To compare time taken, between the testing groups, time must be considered more in detail. In chapter 6.2 on page 66, one hypothesis was made that female users analyze the hotel options more deeply than male users. They care about beauty, location etc., while male users, on average, do the booking process in a straight forward manner without spending time on additional options.

In chapter 8.3.1 observations were made about some users who took longer because of problems; they forgot some filters and in the end, they had to do the booking process once again, or users took longer because they were not able to change the language. The purpose at this point is not to analyze the website's usability and weaknesses, which was done in in chapter 8.3.1 and 8.3.2. Here, the analysis focuses on whether different user groups behave differently during an online booking process. Any users who encountered problems during the booking process, and took time solving these problems, will be excluded from this part, so as not to manipulate the statistics. Since critical users suffered at least one problem, all critical users defined in chapter 8.3.1 will be excluded. In addition to that, two other users could have been identified that had problems with the booking process. User 5 had a problem with filtering and user 33 had problems finding the language button, which increased the booking process time. This means 6 critical users and 2 additional users will be excluded from this analysis, leaving 26 users that will be considered at this point.

Time by Continent

Table 19 – Time by continent

Time by Continent [seconds]

Continent	Max. Time	Min. Time	Avarage Time	Total
Europe	996	278	486	11
Latin America	923	304	482	15
Total				**26**

According to table 19, the maximum time taken by a European was 996 seconds and the minimum time was 278 seconds. The slowest Latin American needed 923 second to book and cancel the hotel and the fastest was 304 seconds. The European average was 486 seconds and the Latin American average was 482 seconds.

The slowest European was a German woman who considered the hotel details very carefully. She checked the hotel's location, read reviews and checked the cancellation agreements. The fastest European did his job in a straight forward manner and booked the hotel without spending a lot of time on optional details. The slowest Latin American did not hurry up and was leisurely about making the booking. In addition, she checked the beauty of the hotels looking at some pictures. Even the fastest Latin American considered more options and opened more tabs with different hotels, he skipped the details and did not go in detail for every hotel.

Time by Brazilian vs. Exchange students

Table 20 – Time by Brazil vs. European and Latin American exchange students

Time by Brazil vs. European and Latin American exchange students [seconds]

Brazil or Exchange	Max. Time	Min. Time	Avarage Time	Total
Brazil	923	304	486	9
Europe	996	278	486	11
Latin America	679	353	476	6
Total				**26**

The average time of all three groups are similar (Brazilian and European average is 486 seconds and the Latin American average is 476 seconds). The other key facts can be seen in table 20 above.

Time by Age Group

Table 21 – Time by age group

Time by Age Group [seconds]

Age	Max. Time	Min. Time	Avarage Time	Total
<20 years old	576	386	481	2
20 - 22	679	337	471	11
>22 years old	996	278	495	13
Total				**26**

Since only two teenagers were part of this group (see table 21), it will be hard to draw any conclusion. The fastest teenager took 386 seconds to complete the booking process and the slowest one took 576 seconds. The slowest person over 22 years old was the German women with 996 seconds mentioned above. The fastest in this age range was the same European as mentioned above with 278 seconds. The slowest person between 20 and 22 years took 679

seconds, once again the slowest Latin American above. The fastest person in this age range was user 17 who completed the task in 337 seconds. He did the process quickly and very efficiently without spending time on optional details.

The average time between the persons aged between 20 and 22 years and older than 22 years was similar. In the younger age range the average was 471 seconds and the older one was 495 seconds.

Time by Gender

Table 22 – Time by gender

Time by Gender [seconds]

Gender	Max. Time	Min. Time	Avarage Time	Total
Female	996	304	553	8
Male	923	278	453	18
Total				**26**

In the following the results of table 22 will be discussed.

The longest time by a female tester was the German student that was already mentioned above with 996 seconds. The fastest female tester was user 8 with 304 seconds, even though she looked at different options and did not choose the first option. The fastest man was user 11 with 278 seconds and the slowest was user 26 with 923 seconds. He checked the cancellation policies as well as checking different alternatives. The average time also shows a difference in this case. The women's average is 553 seconds while the men's average is 453 seconds.

Windows

For this category, the same requirements as for the time category were chosen because the behavior during the booking process can be only evaluated if all users completed the task with the same conditions, which are not failing or no problems occurring. Consequently, a user that had a problem at some point during the process might have to go back to the beginning or go back to the previous window which would influences the statistics. In the following, all tables will be presented at the same time. Afterwards, they will be discussed according to each testing group.

Table 23 – Windows switched statistics for task 1

Windows Switched Task 1 by Continent

Windows Switched task 1	6	7	8	9	10	11	12	13	Total user	Avarage
Europe	3	3	2	0	1	1	0	1	11	8,09
Latin America	2	4	5	2	1	1	0	0	15	7,93
Total	**5**	**7**	**7**	**2**	**2**	**2**	**0**	**1**	**26**	

Windows Switched Task 1 by Brazilian vs. Exchange

Windows Switched task 1	6	7	8	9	10	11	12	13	Total user	Avarage
Brazilian	1	1	3	2	1	1	0	0	9	8,44
European exchange	3	3	2	0	1	1	0	1	11	8,09
Latin American exchange	1	3	2	0	0	0	0	0	6	7,17
Total	**5**	**7**	**7**	**2**	**2**	**2**	**0**	**1**	**26**	

Windows Switched Task 1 by Age Group

Windows Switched task 1	6	7	8	9	10	11	12	13	Total user	Avarage
<20 years old	0	0	2	0	0	0	0	0	2	8,00
20 - 22	1	4	3	1	2	0	0	0	11	7,91
>22 years old	4	3	2	1	0	2	0	1	13	8,08
Total	**5**	**7**	**7**	**2**	**2**	**2**	**0**	**1**	**26**	

Windows Switched Task 1 by Gender

Windows Switched task 1	6	7	8	9	10	11	12	13	Total user	Avarage
Female	0	2	3	0	1	1	0	1	8	9,00
Male	5	5	4	2	1	1	0	0	18	7,56
Total	**5**	**7**	**7**	**2**	**2**	**2**	**0**	**1**	**26**	

Considering table 23, testers that completed the first task without any challenges or problems changed windows between 6 and 13 times.

Half of all European (6 out of 12) opened 6 or 7 windows which are the most efficient performances in this category. And only 2 Latin Americans needed more than 9 windows. The average of Europeans was 8.09 windows and of Latin Americans 7,93.

The Brazilians and European exchange students' division brings a similar average of 8.44 windows for Brazilians and 8.09 for Europeans. Latin Americans opened fewer windows with an average of 7.17 windows. Almost 70% of all users did not open more than 9 windows.

The age diversion is also similar. All averages between all age groups are between 7,9 and 8.1 windows. Remember that only 2 teenagers are considered in this statistic and interestingly, only users over 22 years switched windows more than 10 times.

There is a little more difference between the genders. On the one hand, no women needed less than 7 windows but they opened between 7 and 13. On the other hand, only 2 men out of 18 needed more than 9 windows and most of them needed either 6 or 7 windows (5 users each). This is also reflected in the average which is 9.00 windows for female and 7.56 windows for male.

Mouse Movement

Since the relative mouse movement (mouse movement divided by time) will be considered for this statistic, restrictions used before, for the time or windows analysis, are not necessary. Even if a user had problems and took longer and consequently also might need to move the mouse more often, it will be given in relative terms which means every single user can be considered for this criterion. In the further tables Mouse Movement divided by time will be shortened by MM/Time.

Mouse Movement by Continent

Table 24 – Relative mouse movement by continent

Mouse Movement by Continent

Continent	Max. MM/Time	Min. MM/Time	Av. MM/Time	Total
Europe	246,65	24,49	114,71	13
Latin America	217,56	10,02	115,57	21
Total				**34**

In general, in table 24 the mouse movement does not vary a lot between the tested Europeans and Latin Americans. The maximum relative mouse movement for the European group was 246.65 pixels/s and for the Latin American group was 217.56 pixels/s. The smallest relative mouse movement for the Europeans was 24.49 pixels/s and for the Latin Americans 10.02 pixels/s. Therefore, the average is very similar with 114.71 pixels/s for Europeans and 115.57 pixels/s for Latin Americans.

Mouse Movement by Brazilian vs. Exchange students

In table 25, the average relative mouse movements of each group is similar. However, Latin American exchange students have by far the lowest maximum relative mouse movement (180.79 pixels/s), but on the other hand the highest minimum mouse movement with 63.30 pixels/s.

Table 25 – Relative mouse movement by Brazilians vs. European and Latin American

Mouse Movement by Brazilian vs. European and Latin American Exchange students

Brazil vs. Exchange	Max. MM/Time	Min. MM/Time	Av. MM/Time	Total
Brazil	217,56	10,02	115,10	12
Europe	246,65	24,49	114,71	13
Latin America	180,79	63,30	116,19	9
Total				**34**

Mouse Movement by Age Group

Table 26 – Relative mouse movement by age group

Mouse Movement by Age Group

Age Group	Max. MM/Time	Min. MM/Time	Av. MM/Time	Total
<20 years old	178,47	86,53	125,76	5
20 - 22	246,65	63,27	119,73	13
>22 years old	217,56	10,02	115,24	16
Total				**34**

Taking a look on table 26, it can be seen that the oldest age group has the lowest average with 115.24 pixels/s and, also the lowest minimum mouse movement with 10.02 pixels/s. Only the teenagers can beat this age group with the lowest maximum mouse movement with 178.47 pixels/s.

Mouse Movement by Gender

Table 27 – Relative mouse movement by gender

Mouse Movement by Gender

Gender	Max. MM/Time	Min. MM/Time	Av. MM/Time	Total
Female	246,65	63,27	139,03	12
Male	180,79	10,02	102,26	22
Total				**34**

The women with the maximum mouse movement per second was user 33 with 246.65 pixels/s. Even the second highest female mouse movement per second was close to this user with 217.56 pixels/s. In contrast the highest mouse movement per second in the male testing group was 180.70 pixels/s and the minimum mouse movement per second reached by a man was 10.02 pixels/s. The minimum mouse movement reached by a woman was 63.27 pixels/s. The average also displays a difference. Men reached an average of 102.26 and women of 139.03 pixels/s.

Keyboard hits versus Mouse Clicks

As already mentioned before, it is difficult to consider keyboard hits and mouse clicks in isolation because they may be connected to each other. Some users may prefer to scroll down with the mouse and others with the keyboard. Some switch from one text box to the other with the tab button on the keyboard while others click on the next text box with the mouse.

To get enough data to compare each testing group, two graphs will be shown. One graph will show the absolute keyboard hits and mouse clicks and the other graph will show the relative keyboard hits and mouse clicks per second (keyboard hits and mouse clicks for task 1 divided by time for task 1). In the second step, it will be analyzed if specific testing groups appear in specific areas and if any conclusions can be drawn.

Keyboard hits versus Mouse Clicks by Continent

The following figure 33 shows the relation between absolute mouse clicks and keyboard hits for task 1. The green stars represent European users and the blue points Latin American users. The same goes for figure 34 which shows this statistic relatively divided by the time for task 1. The horizontal axis always represents the keyboards hits (absolute keyboard hits or keyboard hits per second). The vertical axis represents the mouse clicks (absolute mouse clicks or mouse clicks per second).

Figure 33 – Keyboard hits vs. mouse clicks by continent

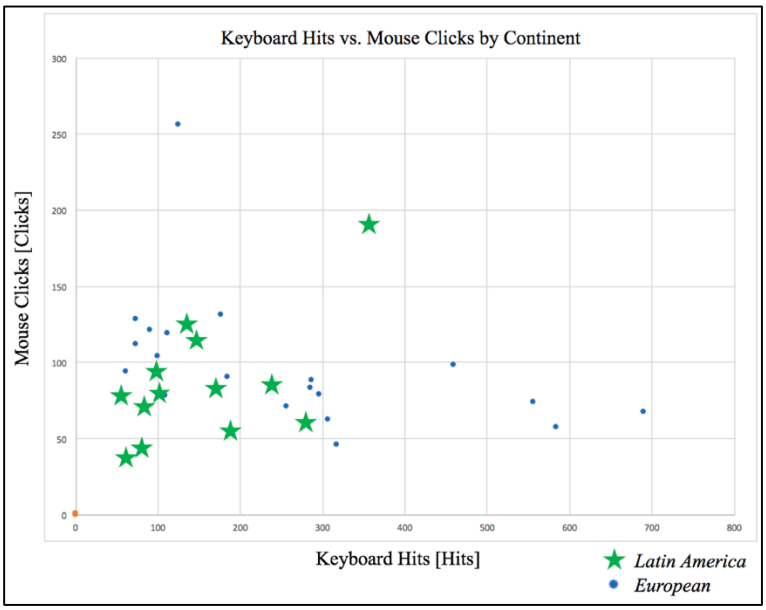
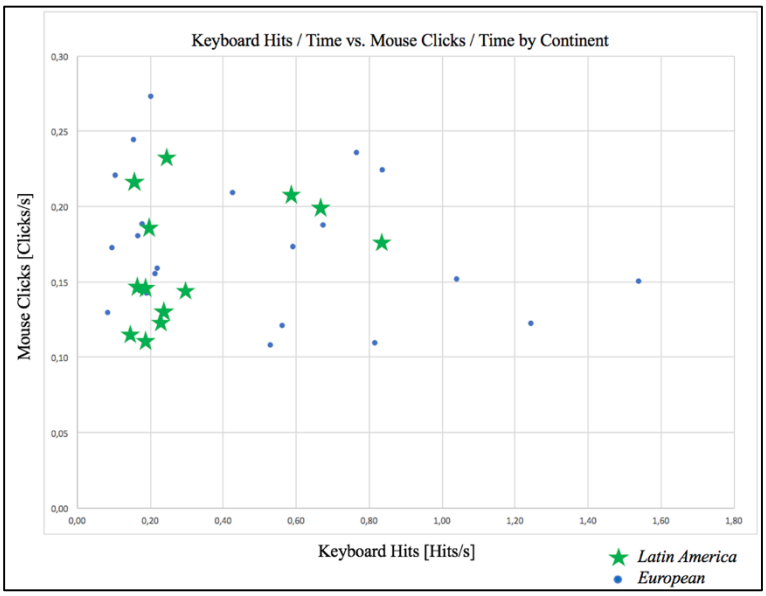


Figure 34 – Relative keyboard hits vs. mouse clicks by continent



Keyboard hits versus Mouse Clicks by Brazilians vs. exchange students

The following two graphs (figure 35 and 36) show the division between Brazilian students and exchange students. In these graphs the blue points represent the European users, the orange squares the Brazilian testers and the green stars the Latin Americans.

Figure 35 – Keyboard hits vs. mouse clicks by Brazilian vs. exchange student

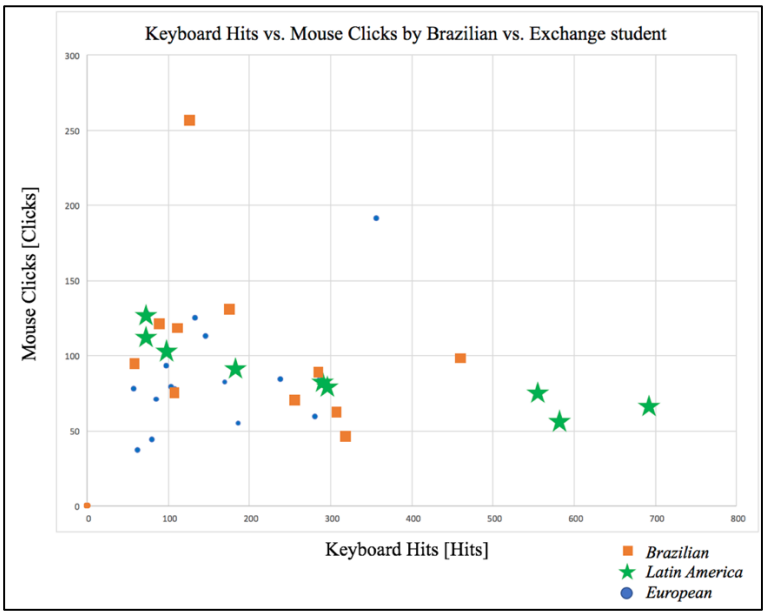
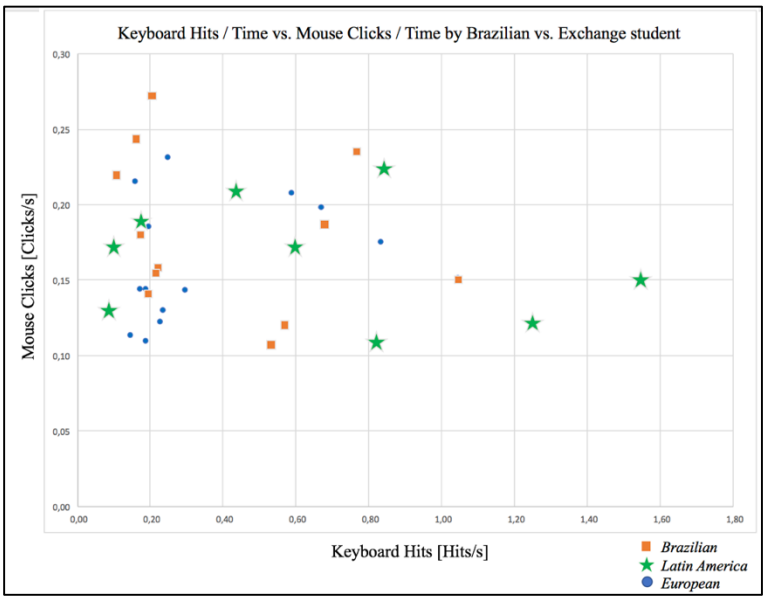


Figure 36 – Relative keyboard hits vs. mouse clicks by Brazilian vs. exchange student



Keyboard versus Mouse Clicks by Age Group

In the following graphs (absolute keyboard hits versus mouse clicks in figure 37 and relative keyboard hits versus mouse clicks divided by time for task 1 in figure 38) the yellow squares represent teenagers; the purple stars the users between 20 and 22 years and the blue points all users older than 22 years.

Figure 37 – Keyboard hits vs. mouse clicks by age group

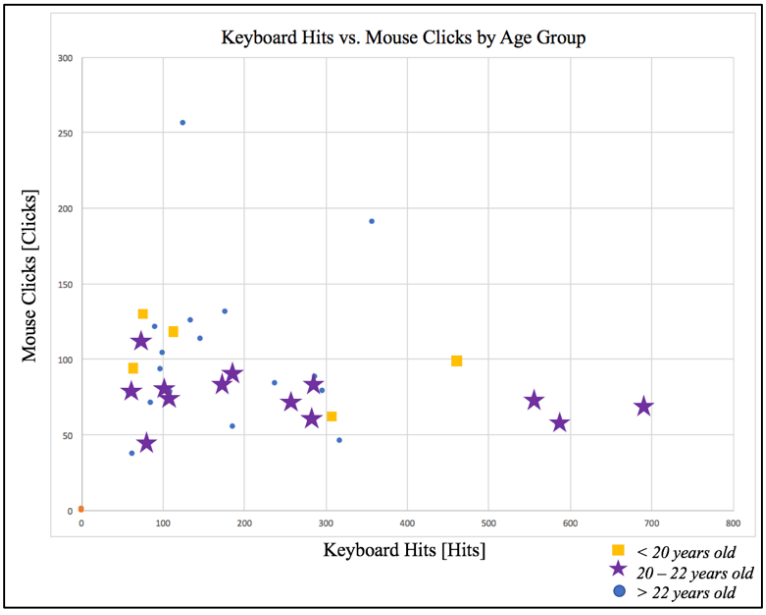
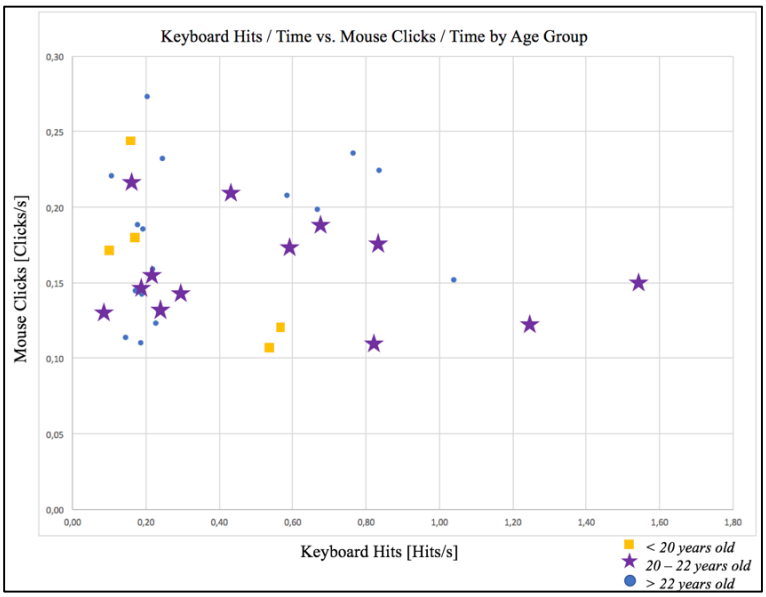


Figure 38 – Relative keyboard hits vs. mouse clicks by age group



Keyboard versus Mouse Clicks by Culture

The last two figures 39 and 40 repeat the whole process only with gender division. The red stars represent female users and the blue points male users.

Figure 39 – Keyboard hits vs. mouse clicks by age gender

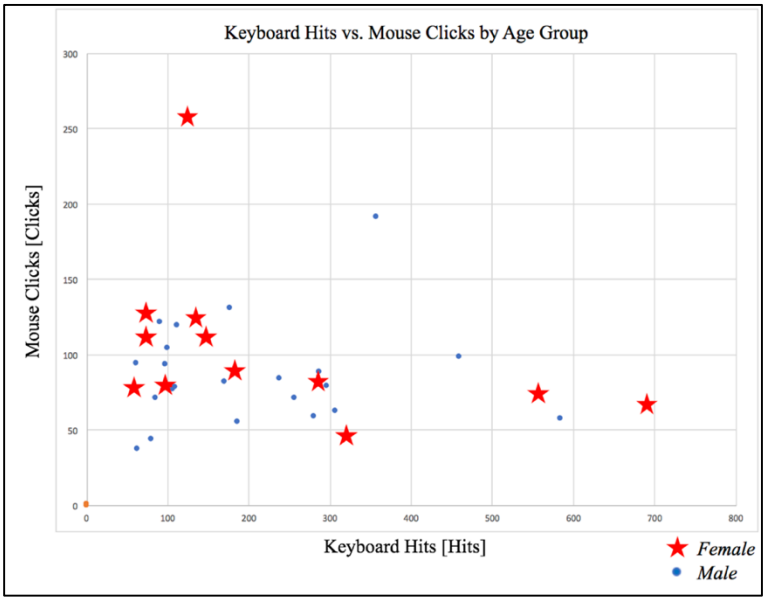
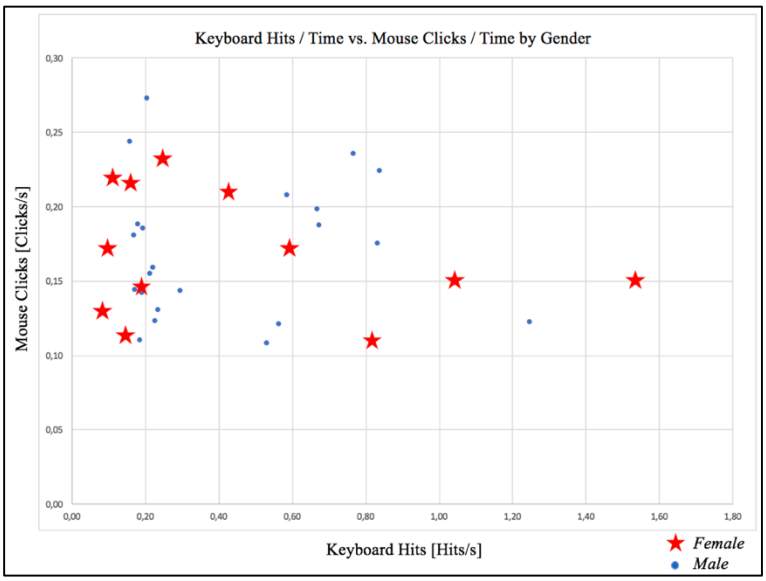


Figure 40 – Relative keyboard hits vs. mouse clicks by gender



8.4 SUMMARY

This chapter delivered a detailed explanation of the testing procedure. First, the MORAE© software and the Excel file were presented and their valuable support to this thesis. After this, the testers were presented and divided into testing groups regarding country, exchange students, age range and gender. This was followed by all results and statistics revealed by the collected data.

In the next chapter these results will be discussed and conclusion will be drawn.

9 DISCUSSIONS & RECOMMENDATIONS

In the previous chapter, the results were presented in forms of numbers. Key users and key performances have been looked at. In this chapter these results will be interpreted; recommendations will be given for the website and conclusions about UX factors will be drawn. And finally, the hypotheses made in chapter 6.2 will be classified. Before entering the results that were collected from the test, the method will be validated.

9.1 VALIDATION OF METHOD

This sub-chapter will validate the developed method in relation to the requirements defined in chapter 5.1 on page 59. Following requirements were defined: usable, practicable, flexible, transparent, reproducibility and complementary.

Since this method has given useful and logical recommendations, the method meets the requirement of being useful. The exemplary application of this method did not exceed financial resources and could be realized in 4 days, which confirms to the practicable requirement. In this case, the hotel booking website Booking.com was evaluated. However, any other website can be tested on the same way which makes the method flexible. Anyway, this method is not even limited on websites. It can be used for any other technological usage which might be related to the Internet of Things (IoT) such as smart devices or Smart TV. Last but not least, as every step was deeply explained and all results were illustrated by tables and all software were also explained by illustrating screenshots, the method was transparent and reproducibility at all time.

9.2 TASK 1

In chapter 8.3.1. key users were defined. At this point, a special focus will be given on the critical users. Keep in mind that user 14, 16, 18, 19, 29, 34 were critical users and each of them showed up at least one problem during the booking process (see table 8 on page 99). Following problems can be summarized and in brackets the number of users suffered this problem are listed:

- Failed the first time (4; User 16, 18, 29, 34)
- Problems with finding filters or did not use filters (5; User 14, 16, 18, 29 34)
- Did not have control over the process (5; User 14, 16, 18, 19, 34)

- Problems with switching language (2; User 19, 34)

Obvious problems of the critical users were finding the filters (5 out of 6 persons had problems with filtering) and not having the control in the process (5 out of 6 persons seemed to be lost in the process and did not have any control in the process). These two urgent problems are chosen to be analyzed more in detail. The other problems are suggested to be analyzed in further steps since time limitation of this research does not allow to do a deep analyze of all problems.

As already mentioned in chapter 6.4 on page 74, the 5-Why method will be used to discover the origin of the problem. In the following the first problem, finding the filters, will be analyzed which is followed by the second problem, no control in process.

Finding the filter

The first why will be answered by the first criterion of Bastien and Scapin, which is "Guidcance; Grouping/Distinction of Items". The answer of asking why this criterion failed is the second sub criterion "Grouping/Distinction by Format". Analyzing this problem, it can be explained by poor visual distinction of data. Booking.com provides a text box that calls attention to the filter. This appears as soon as the user gets to the hotel options. Going once again deeper into the reason, it can be concluded that the reason of failing was the visually unfavorable white text box which was put on a white background makeing it hard to read. At some users, this text box did not even appear which means Booking.com does not even call attention to their filters.

No control in process:

Having no control in the process is an obvious indication of Bastien & Scapin's third criterion "Explicit Control". Going deeper, it targets the "User Control" which is the second sub-criterion of "Explicit Control". If an interface provides a good user control, the interface becomes predictable and it diminishes the probability of making errors. However, this was not the case because the users, that suffered this problem, were not able to predict the website at any time of the process. The reason of this is that they were not able to accomplish the task and were overwhelmed by the website. This leads to the same reason than before, because the user mostly did not find the filter that was supposed to ease the process and predict the further steps of the booking.

Now, the critical users' auto evaluation will be analyzed. Questions of the auto evaluation that were rated with the worst rating by each critical user will be listed. Keep in mind questions 6 to 20 relate to Bastien and Scapin (1993) usability criteria. Each question targets one specific criteria. If one question gets many weak evaluations, this criterion will be analyzed in detail. In this case question 4, 13, 15, 20 and all error management questions (16, 17, 18) got the overall worst ratings. The list of each question that got the worst or best rating by each critical user, can be seen in appendix D.

Looking at table 6 in chapter 6.3.1 on page 69, question 13 targets Bastien and Scapin's "Workload and Minimal Action – criteria". Question 15 targets the website's Grouping/Distinction of Items. Question 20 the predictability and expectations of the website (Explicit Control) and question 4, which was rated badly as well, claims the poor satisfaction and it also targets parts of the User Control criterion. The poor "Grouping/Distinction of Items" was already analyzed in problem 1 (Finding the filters). The poor predictability was deeply analyzed in problem 2 (No control in process) and the bad "workload and minimum action" criterion and the poor satisfaction will be added to the Ishikawa Diagram but will not be deeply analyzed at this point. The questions that targeted the error messages will be not considered at this point because some comments about error messages will be done in the further research.

In the following an Ishikawa Diagram (figure 41) will be presented with the problems mentioned above. The bold problems indicate the urgent problems that were analyzed. To not lose the track, only two dimensions of the 5-Why method will be included into the Ishikawa Diagram and all other why-questions will be summarized in table 28.

Figure 41 – Ishikawa diagram for task 1

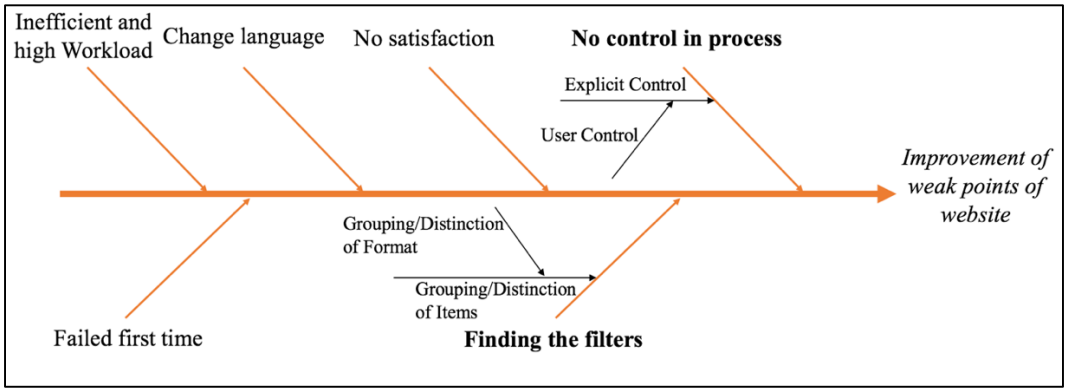


Table 28 – 5-Why method for problems task 1

Problem 1: Finding the filters		**Problem 2: No control in process**	
Why-Question	**Answer**	**Why-Question**	**Answer**
Why did not the user find the filters?	Lack of Bastien and Scapin's first criterion Grouping/Distinction of Items.	Why did the user have no control in the process?	The third criterion of Bastien and Scapin was not adequate (Explicit Control).
Why was there a lack of Grouping/Distinction of Items?	The sub-criterion Grouping/Distinction of Format was not sufficient.	Why was not Explicit Control adequate?	There was a lack of the second sub-criterion, User Control.
Why was not the Grouping/Distinction of Format not sufficient?	Poor visible distinction of data.	Why was there a lack of User Control?	Booking.com did not provide predictability of the website.
Why was there a poor visible distinction of data?	White text box over white background indicated the filter function which was poorly identifiable.	Why did not Booking provide predictability?	The user was overwhelmed during the process.
		Why was the user overwhelmed during the process?	The user was searching frustratingly for ad equated hotels which led to the lack of finding the filters again.

The following two recommendations will be given to the website:

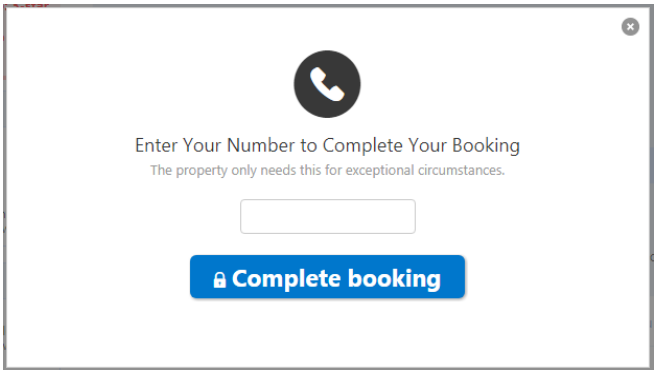
- Guidance and Grouping/Distinction of Items should be improved, considering the analyzes above. Since the problem was at finding the filters, a clearer visual distinction of data could be provided. Booking.com provides a text box that calls attention to the filter. This appears as soon as the user gets to the hotel options. However, this is visually unfavorable because the white text box is put on a white background which makes it hard to read. An example of an improvement could be a pop-up window that appears when opening the interface with the filters. This pop-up window would draw attention to the filters and to the basics of using the website. Everybody pays attention to pop-up windows because they must be actively clicked away to disappear. Users that are familiar with the process and with filters may have the option to check a "never ask again" checkbox which would prevent the pop-up window appearing again and annoying the experienced user.

- Another suggestion, that targets the Explicit Control – User Control, can be a live chat. Some users felt lost during the process. A live chat makes it possible to eliminate doubts and questions immediately. One user commented that he prefers to book with agencies because he has a person to talk to if he has any doubts. Only the presence of someone talking to the user would probably help. And in case of any problems and questions, the person can help immediately.

However, besides those weak points, some mentionable and admirable features should be pointed out. As already mentioned, 5 out of 6 critical users had problems with finding the filters. Although, all users figured out, sooner or later, how to use the filter option. 3 out of these 6 critical users rated question 7, that targeted the filter feature, with their highest rating. This means, even though problems occurred by finding the right filters, after resolving this, the filter worked nicely and stood out in the user's mind.

Another point worth mentioning is the error message. Question 16, 17 and 18 target the error management. Even though 3 critical users rated those three questions as the worst, 7 out of the 34 users clicked the options that no error message appeared. However, 6 out of these cases were confronted with error messages. Some users forgot to fill out the name, others forgot to fill out the phone number. But Booking.com does not show that the user did an error and forgot to fill out a box. When the user wanted to continue, Booking.com scrolled up automatically to the missing box and let the box description appear in red. Or if the phone number was missing, Booking.com did not even scroll back to the blank box, instead opening a pop up window which asked for the phone number, meaning the user did not even realize he had forgotten to put in this information. Figure 42 demonstrates this smart method.

Figure 42 – Phone number entry



9.3 TASK 2

Considering the results shown in chapter 8.3.2, more than 70% did not complete this task and most of them ended up being angry and depressed. The MOM is 14 seconds for this task and even the best performer took 81 seconds. In fact, he is the only user that took less than 3 minutes. It is obvious that contacting the customer support is a big challenge in this case.

Applying the 5-Why method to this problem, the first question can be answered by the first criterion of Bastien and Scapin, missing “Guidance; Grouping/Distinction of Items”. Going deeper into the problem, it can be observed that it is the problem of the first and second sub-criterion “Grouping/Distinction of Location and Format”. The path and the buttons that lead to the customer service box must be more visible and more easily identifiable. At the moment there are not. Obvious buttons like “contact us” do not lead to the goal. In contrast, only nondescript buttons like “Still need help?” or “Others”, very small written, lead to the customer service box. To improve this problem a better description, better location, increase of text size, bold text or more obvious buttons should be used.

In addition to that, the third criterion of Bastien and Scapin, “Explicit Control”, also suffered at this point. More specific, it was the “User Control” (second sup-criterion). Frustration of the user was observed, which was ended up in giving up. The main reason why most user gave up was because they obviously did not find the customer box in an adequate time range. This can be reasoned by the same critical points than shown in the paragraph above (no obvious buttons that lead to the customer service box, unfavorable text size and colors etc.).

This main problem can be seen in figure 43 which illustrates the previous Ishikawa Diagram that was created for the previous task 1. The only difference to the previous diagram is that this diagram includes the problem for task 2 including two dimensions of the 5-Why method. The diagram will be followed by table 29 with the full 5-Why analyze for this specific problem.

Figure 43 – Ishikawa diagram for task 1 and task 2

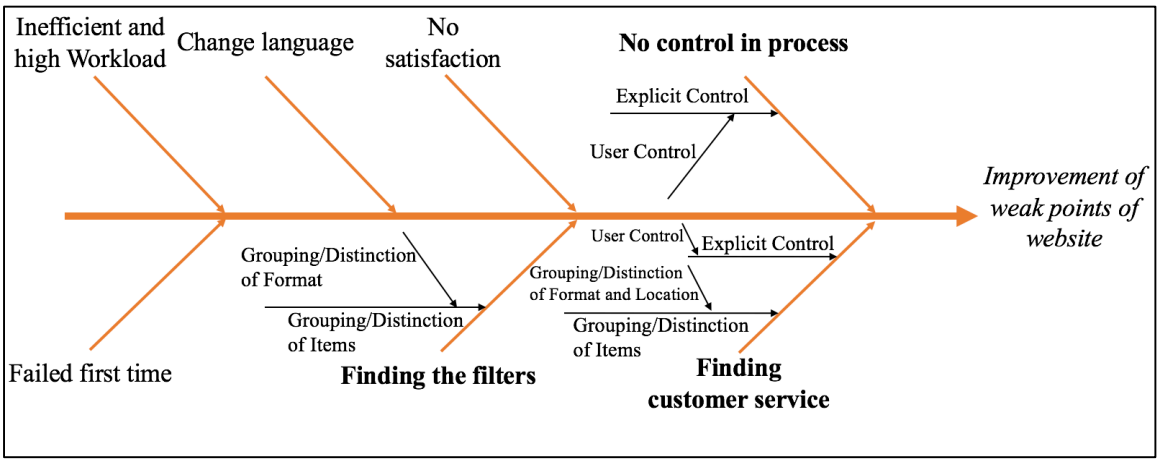


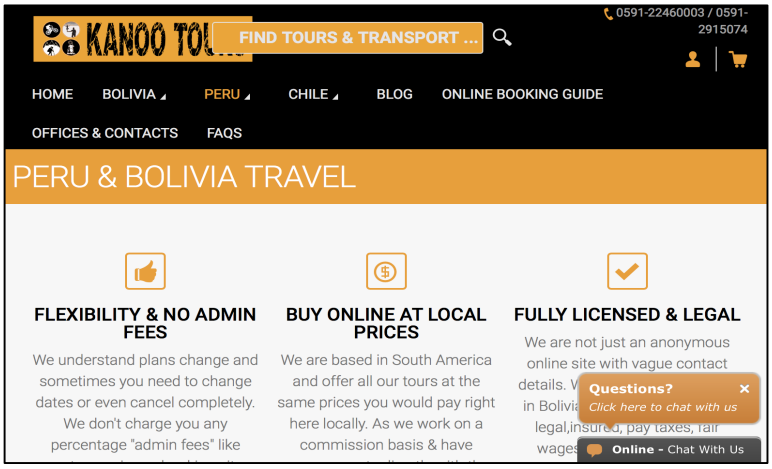
Table 29 – 5-Why method for problem task 2

Problem: Finding the customer service	
Why-Question	**Answer**
Why was it challenging to find the customer service box?	Lack of Bastien and Scapin's criteria Grouping/Distinction of Items and Explicit Control.
Why was there a lack of Grouping/Distinction of Items and of Explicit Control?	The sub-criteria Grouping/Distinction of Format and Location was not sufficient, as well as User Control.
Why was not the Grouping/Distinction of Format and Location as well as User Control not sufficient?	The description of the buttons to open the customer service box were not clear and visibility of buttons were not easily identifiable (nondescript, small size). User was not willing to continue the process and gave up.
Why did user give up?	User did not find customer service box in an adequate time.
Why did not the user find the customer service box in an adequate time?	The description of the buttons to open the customer service box were not clear and visibility of buttons were not easily identifiable. (nondescript, small size).

An interesting recommendation for this problem would be the live chat that has been already recommended for the previous sub-chapter. In fact, in this case, it would improve the customer service massively. On the one hand, the user would not have to search for the customer

service box anymore, and on the other hand, the user would receive his answer immediately and would not have to wait for the reply. Many travel agencies have a live chat that appears on every page on the button right of the website. In figure 44 an exemplary live chat window of the Kanoo Tours traveling agency is illustrated (KANOO TOURS, 2017).

Figure 44 – Live chat example (KANOO TOURS, 2017)



Nevertheless, maybe Booking.com does not want users to send messages to the customer service because this would mean more employees needed to take care of the messages. So, they make it hard to find the direct message conversation to avoid too many messages. Clearly, a good internet platform and service also needs a good customer service which is available for its clients.

9.4 UX FACTORS

Culture

Comparing Europe with Latin America, an obvious difference between the extreme points can be observed on page 101. Europeans lead the best performances. Almost 2/3 of the best performances were Europeans and more than 80% of all critical users that were defined came from Latin America. However, this big difference can be explained by the online booking experience. Europeans are far more familiar with online booking and even with the specific website Booking.com than Latin Americans. While every European books at least one hotel online a year, 19% of Latin Americans never book hotels online. It goes even further. 60% of all Latin Americans book no more than one hotel a year. In contrast, only 15.38% of European do not book more than 1 hotel a year.

Considering time, it is hard to draw any conclusion. On the one hand, there are European users that had no problems during the booking process and considered different hotel options during the booking process. They took their time, did not hurry and booked, in their opinion, the best possible option. On the other hand, there were European users that handled the process in a straight forward manner, only considered the obligatory constraints and did not spend much time on optional features. In contrast, both the slowest and the quickest Latin Americans at least compared the chosen hotel with other options. However, if Latin Americans tended to analyze more options, the switched windows would confirm this. Since different windows must be opened to consider different options, in this case Latin Americans would switch the windows more often. However, this is not the case. The average European opens 8.09 windows per booking process. The average Latin American opens 7.93 windows (compare on page 111). This does not fit in with the assumption made above which therefore does not verify the hypothesis that Latin American book more calmly and tend to analyze more hotel options.

There is also no significant difference between the mouse movement behavior. Europeans maximum relative mouse movement was 246.65 pixels/s and the Latin American 217.56 pixels/s. The minimum mouse movement differs a little bit (10.02 pixels/s for Latin Americans and 24.39 pixels/s for Europeans). However, the average mouse movement reached by both is very similar with 114.71 pixels/s for Europeans and 115.57 pixels/s for Latin Americans (compare page 112).

Considering the keyboard hits vs. mouse clicks behavior, according to the graphs in 8.3.3 page 114 and 115, it seems that Europeans tend to use more the keyboard while the Latin Americans tend to favor the mouse. While almost no Latin American represents the right part of those graphs, some European can be found on this part in both graphs (the absolute graph of figure 33 and the relative graph of figure 34). The right part of the graph means using frequently the keyboard. However, it also seems that many Europeans prefer to use the mouse which is illustrated by many European users on the left side which is means less keyboard and more mouse clicks. All in all, it cannot be concluded in general that Europeans prefer using the keyboard.

Remembering the hypothesis made above, exchange students may have more experience than Brazilians because they are out of town, do not live with their parents and want to explore the country during their exchange. As already observed above, the division of exchange students

and Brazilian students would not be fair enough. Another division must be done between the exchange students (European and Latin American exchange students). It will be quickly clear that the hypothesis cannot be verified that easily. Most rewarded performances came from European exchange students (66.67% of all best performers were European, see page 102). Only one Latin American and only one Brazilian were rewarded as best performer. Furthermore, half of all critical performers were Latin American. Every third was Brazilian and only 1 out of 6 critical persons was European (16.67%). This can again be logically put down to experience. Almost every second European exchange student books at least 7 hotels online a year, while none of the Latin Americans do so and only 8.33% of Brazilians. In addition, every European books at least one hotel online while more than 60% of Brazilians and Latin Americans book maximum 1 hotel a year.

Furthermore, every European books hotels online (page 105), when he needs a hotel. Most Latin Americans and Brazilians also do, but not all of them (66.67% of Brazilians do and 77.78% of Latin Americans). Interesting is also that every European books on Booking.com while only every third Latin American does. Even Brazilians book more on Booking.com with 66.67%.

In conclusion, Europeans are much more experienced in online booking processes than Latin Americans and Booking.com is more aimed at Europeans than at Latin Americans. Other conclusions, such as Latin Americans may book with more calm or analyze more deeply than Europeans could not be verified. Moreover, it was observed that more Europeans used the keyboard more frequently to go through the booking process than Latin Americans. However, to draw any conclusion out of this, more tests are required.

However, obvious is that European's are more experienced than Latin Americans. Of course, the reason can be that Europeans travel more than Latin Americans, but Booking.com can take an expansion as a consideration. Better marketing can lead to more customers in those regions.

Age

Age was divided into three testing groups, teenagers (less than 20 years old), people between 20 and 22 years old and people older than 22 years old. It will be quickly clear that no

teenager is part of the best performance group while 40% of all teenagers were part of the critical users. In the age group between 20 and 22 years 7.69% were best performers and in the oldest age group more almost every third was grouped as one. Consequently, no teenager is part of the best performance, while 16.67% were between 20 and 22 and 83.33% were older than 22 years. Considering the critical performance, half of all critical performers were older than 22 years old. This means the age group older than 22 years is greatly represented among, on the one hand, the best performers, and on the other hand, the critical performers. This can be explained by the high number of participants in this age group. These statistics can be seen on page 103.

Considering experience on page 106, a couple of conclusions can be drawn. Teenagers are poorly experienced. 60% of all teenagers do not book hotels more than once a year and 80% not more than twice a year. In contrast, half of all people older than 22 years book at least 2 hotels a year indeed every third person even books more than 6 hotels a year. In the middle age group, 20 to 22 years, around 50% books more than 1 hotels a year. Moreover, 87,50% of the older age group uses Booking.com. In the middle age group, 53.85% uses Booking.com. The two older age groups also book hotels, when they need to book one, online, with more than 80%. Whereas barely 60% of all teenagers claimed that. While every second person of the older age group considers the internet as a place for online booking, no teenager does.

Considering these numbers, the online booking experience increases with the age. Even though fewer teenagers were asked, many of them show poor online booking experience. Considering user comments, young users admitted that they let their parents book because many of them still lived with their parents and travel with them. This may change with age.

Relating teenagers to time is impossible, since only two teenagers managed the tasks. And comparing the time differences between the other two age groups, no big differences can be seen. In both age groups, there are people that completed the process in a straight forward manner and others analyzed alternatives. The average between those two age groups is a difference of only 5% (471 seconds and 495 seconds for older than 22 years, compare to table 21 on page 109).

The mouse movement of the oldest age groups seems to be most efficient (compare to page 113). Both the maximum mouse movement per second, and the minimum mouse movement per second is by far less than the middle age group. Even the three lowest mouse movements are

lower than the minimum mouse movement of the middle age group. The average also tends to be less. If this is considered as a difference, which should be considered carefully, this can be put down to experience. Since older people are more experienced they may move the mouse more quickly to the right spot.

Having a look at the graphs of figure 37 and figure 38 on page 116 and 117, some exceptions on both graphs (relative and absolute graph) in the age group older than 22 can be identified in the right part of the graph. This means they used in comparison to the others a lot more the keyboard. Both older age groups (between 20 and 22 years and older than 22 years) can be also found in the middle of both graphs which means they used a mix between mouse and keyboard. However, most users in all age groups, still are represented in the left side of the graph which means they all used more the mouse in comparison to the keyboard. Again, no obvious results can be observed which does not lead to any conclusion.

Summarizing, it is difficult to draw any conclusions about teenagers because only 5 teenagers were asked in the whole experiment. However, it seems that the older the user, the more experienced he is. Besides that, no other conclusion could have been observed.

Gender

Considering the extreme performers on page 103, exact 50% of all critical users were male and the other half were female. However, every fourth woman was part of the critical group and barely 13% of all men were part of this group. Having a look at the best performers, none of all women was part of the best performer consequently 100% were male. In other words 27,27% of all men tested were best performer.

So, having a deeper look, even if it seems to be similar in some parts, it must be mentioned that almost twice as many men were tested (12 women to 22 men). These numbers showed that in general women performed worse than men. It cannot be concluded that experience is a reason of this performance difference. There is no mentionable difference between the booking experience on table 18 on page 107. Anyway, to conclude that men are more used to online booking websites than women, may be a reason, but this thesis will not underline this. Instead, this thesis will suggest having a deeper look at this point and doing more tests with men and women.

With the time criterion, a difference can be seen. The women's average is 553 seconds and the men's only 453 seconds (see page 110). Additionally, the maximum time and the minimum time are higher for the female users. Having a deeper look at the user, it will be quickly clear that most of the women paid attention to the beauty of the hotel, location and reviews. Men did it as well but not as often as women did. This observation is strengthened by the number of windows changed statistic. The average women's windows changed was 9 windows, the men's windows changed was 7.56 windows. There were women changing the windows more than 13 times, while only two men changed the windows more than 9 times. No women completed the task without changing more than 6 times, while 5 men out of 18 changed the window only 6 times (page 111).

Mouse movement is also interesting to analyze. The maximum mouse movement per time and minimum mouse movement per time, as well as the average varies a lot. In all categories men used the mouse significantly less. Now, analyzing some key users, it is clear that some women moved the mouse without paying attention. Two of them even moved the mouse, while analyzing the website, to wherever they were looking. Other women draw circles on the window without having specific targets. Men tend to look first and as soon as they know their target they direct the mouse to that certain spot.

Concluding something about the keyboard vs. mouse click behavior is once again difficult because of the bad women/men distribution. However, looking on the absolute keyboard hits versus mouse clicks on figure 39 on page 117, only one women used less than 65 clicks, while 6 male users used less than this amount. This might be an indication of a deeper analyzes of the task. If more options are considered, more pictures are looked at etc. more clicks are needed. This would confirm the statement given above that female user analyze the task more in detail than male user. However, any specific observation about who prefers the mouse and who prefers the keyboard cannot be drawn from these figures.

Summarizing, both men and women seemed to have the same experience regarding online booking. The hypothesis that women care more for beauty and details could be verified by this study. Since this observation is a very powerful one, Booking.com can think of customizing the interface. This means if a female user is using the website, other functions will be highlighted than a male user.

9.5 VALIDATION OF HYPOTHESES

The following table 30 will validate the hypotheses made in chapter 6.2 on page 66 with the results and discussions made in the previous chapter.

Table 30 – Validation of hypotheses

Hypothesis	Validation True or Not	Comments	Recommendation
Booking.com is user-friendly and only few user will fail on task 1.	True	Booking is user friendly, indeed. However, the filter function can be identified easier.	Make the filter function more obvious and more visible, e.g. by a pop-up window.
More user will fail on task 2 but majority will complete successfully.	Not True	More than 70% did not find the customer service.	Be present and available. Let the user enter in contact easily, e.g. a Live Chat.
Latin American will take more time than Europeans.	Not True	No difference in time between Latin Americans and Europeans.	Realize more tests and consider this criterion more in detail.
Exchange students have more experience than Brazilians.	Not True	Latin American Exchange students do not have more experience than Brazilians. However, European Exchange students do.	Booking.com is well known in Europe. In contrast in Latin America not. Booking.com can expand or do a better marketing especially in Latin American countries.
Women analyze the hotel details deeper than men.	True	Women did spend more time on analyzing the hotel details.	Customize the user interface. Different user interface between women and men.
The older the person the more experience he has.	Cannot be validated	Not enough testers to draw credible conclusions.	Need more references and more testers.

9.6 SUMMARY

This chapter discussed the results presented in the previous chapter. Recommendations for the website were given regarding the problems discovered during the analysis. Supported by Ishikawa diagrams and the 5-Why method, the problems and their reasons were illustrated. In the last part of this chapter, conclusions about the differences of UX factor groups were made. Some hypotheses were verified and others could not be, due to lack of information or no significant differences observable.

10 CONCLUSION AND STUDY EXTENSION

This thesis aimed at a usability evaluation for an online hotel booking website which evaluated, on the one hand, the weak points of the website and, on the other hand, differences between different testing groups. These testing groups build different user experience factors that may impact usability during the online booking task. Confirming the research questions in chapter 1.3, hypotheses were made, regarding the UX factors, and were tested.

The first part, the evaluation of the website, was realized by defining criteria and analyzing the worst performances as well as user's feedback. As the criteria, as well as the user's feedback, were based on Bastien and Scapin's (1993) usability criteria. Out of the worst feedbacks and performances, conclusions about weak point of this website could be done. In the second part, testers were split into testing groups. These testing groups were compared to each other and differences were drawn regarding to the hypotheses made before. Some hypotheses could be verified; others were rejected and for others no conclusion could be drawn because the collected data was not sufficient. There were some differences observed between European and Latin Americans, as well as between women and men.

This usability tool can help web designer to identify weak points of the websites. Besides that, web designer can evaluate important UX factor that may have impact on their interface. Even MORAE from TechSmith can think of integrating this method in its software, because so far, this method needs Microsoft Excel to accomplish analysis.

Moreover, further steps are recommended because this research brought some limitations due to time and resource limitation. Mainly the weak distribution of the testing population was challenging. Few teenagers were tested or almost double of men compared to women. Besides that, only a specific group was tested. Social and educational conditions were reduced to a student and higher educational environment. Culture was also limited to continents which is a big limitation because single nations, religions, intercultural influence or even regions were ignored. One person might be born in one place and might grow up in another place or parents come from different countries etc. Further steps might be more tests with a better population distribution and a more complex consideration of the factors. A deeper culture consideration or a more flexible educational environment can be realized that does not only consider higher education.

REFERENCES

ABRAHÃO, J. I., et al. **Ergonomia e Usabilidade - Em ambiente virutal de aprendizagem.** São Paulo: Blucher, 2011.

ABRAHÃO, J. I., et al. **Introdução à ergonomia: da prática à teoria**. São Paulo: Blucher, 2009.

ALLHUTTER, D., et al. **Gender-UseIT: HCI, Usability und UX unter Gendergesichtspunkten**, Walter de Gruyter GmbH & Co KG: Berlin, 2014

AMAZON. Seatle, USA. **Über 1-Click bestellen**. from <https://www.amazon.de/gp/help/customer/display.html?nodeId=201889620>. Retrieved 13.07.2017.

AUSTIN, Z. J. SUTTON. **Qualitative Research: Getting Started**. In: The Canadian Journal of Hospital Pharmacy, Ottawa, Canada, 2014.

BASTIEN, J. C.; D. L. SCAPIN. **Ergonomic criteria for the evaluation of human-computer interfaces**. France: Inria, 1993.

BATH, C. **De-Gendering Informatischer Artefakte: Grundlagen einer Kritisch-Feministischen Tedchnikgestaltung**. 2013 Dissertation - Mathematik & Informatik. Universität Bremen, Bremen, 2013.

BAUMANN, K.; H. LANZ. **Mensch-Maschine-Schnittstellen elektronischer Geräte : Leitfaden für Design und Schaltungstechnik**. Berlin: Springer, 1998.

BELL, N. E. Princeton, NJ, USA. **Data Sources: International Comparisons of Educational Attainment.** from <http://cgsnet.org/data-sources-international-comparisons-educational-attainment>. Retrieved November 2011.

BICHENO, J. **The New Lean Toolbox: Towards Fast, Flexible Flow**. South Africa, Picsie Books, 2004.

BOOKING.COM. **Booking.com** Amsterdam. from <https://www.booking.com/>. Retrieved November 2017.

BRAJNIK, G.; C. GIACHIN. **Using sketches and storyboards to assess impact of age difference in user experience.** In: INTERNATIONAL JOURNAL OF HUMAN-COMPUTER STUDIES. New Jersey, 2014 p.552-566. v.72.

BROOKE, J. **SUS - A quick and dirty usability scale**. **Usability evaluation in industry**. London: Taylor & Francis**,** 1996. p.189-194.

BUNDESZENTRALE FÜR POLITISCHE BILDUNG. **Frauen in Führungspositionen**. Berlin, Bundeszentrale für politische Bildung, 2010.

CHRISTENSSON, P. **WWW Definition**. from <https://techterms.com/definition/www>. Retrieved November 2017.

CLEMMENSEN, T., et al. **Making HCI theory work: an analysis of the use of activity theory in HCI research.** In: Behaviour & Information Technology, London, 2016. v.35.

COCOTAS, A. **Across Markets, Smartphone Penetration Skews Young.** In: Business Insider Australia, Sydney, 2013.

DIN EN ISO 9241-11. **Ergonomische Anforderungen für Bürotätigkeiten mit Bildschirmgeräten.** Berlin: Beuth, 1998.

ENGESTRÖM, Y., et al. **Perspectives on activity theory.** In: Cambridge University Press, Cambridge: 1999.

FORBES. New York, USA. **Are OTAs Really Killing Brick And Mortar Travel Agencies?** from <https://www.forbes.com/sites/steveolenski/2015/04/27/are-otas-really-killing-brick-and-mortar-travel-agencies/ - 2ffe3f2811d7>. Retrieved October 2017.

FRANDSEN-THORLACIUS, O., et al. **Non-universal usability?: a survey of how usability is understood by Chinese and Danish users**. In: PROCEEDINGS OF THE SIGCHI CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS, ACM. 2009, Boston, MA, USA. 2009.

HALE, K. Stanford, USA. **How to Build Products Users Love.** <https://www.youtube.com/watch?v=2t0qXVkDN>. Retrieved September 2016.

HOFSTEDE, G. Groningen, Netherlands. **Cultural Dimensions.** from <https://geert-hofstede.com/national-culture.html>. Retrieved July 2017.

IRRENHAUSER, T. **Bewertung der Wirtschaftlichkeit von RFID im Wertschöpfungsnetz.** 2014 Dissertation, München iwb, Technische Universität München**,** München, 2014.

ISO 9241-100. **Introduction to standards related to software ergonomics**. Geneva, Switzerland: ISO, 2010.

KANOO TOURS. **Kanoo Tours**. from <https://www.kanootours.com/>. Retrieved October 2017.

KEMPTHORNE, O. **The design and analysis of experiments.** Oxford, England: Willey Eastern Private Limited, 1952.

KIM, H. K., et al. **How User Experience Changes over Time: A Case Study of Social Network Services.** In: Human Factors and Ergonomics in Manufacturing & Service Industries, 2015. p. 659-673. v.25.

LANOUE, S. San Francisco, CA, USA. **What's the diffference between UI and UX?.** from <https://www.usertesting.com/blog/2016/04/27/ui-vs-ux/>. Retrieved 27.04.2016.

LIEBAL, J.; M. EXNER. **Usability für Kids : ein Handbuch zur ergonomischen Gestaltung von Software und Websites für Kinder.** Wiesbaden: Vieweg + Teubner, 2011.

LINDEMANN, U. **Methodische Entwicklung technischer Produkte : Methoden flexibel und situationsgerecht anwenden**. Berlin: Springer, 2009.

LIU, L.; M. T. ÖZSU. **Encyclopedia of database systems**. Berlin: Springer, 2009.

MARSDEN, N.; U. KEMPF. **Gender-UseIT**. Heilbronn, Germany: De Gruyter, 2014.

MWANZA, D.; Y. ENGESTRÖM. **Managing content in E-learning environments.** In: British Journal of Educational Technology, London, 2005. p.453-463. v.36.

NIELSEN, J. CA, USA. **Usability 101: Introduction to Usability**. from <https://www.nngroup.com/articles/usability-101-introduction-to-usability/>. Retrieved 03.02.2017.

NIELSEN, J.; R. MOLICH. **Heuristic evaluation of user interfaces**. In: Proceedings of the SIGCHI conference on Human factors in computing systems, ACM, Denmark, 1990.

NORMAN, D. **The term "UX". The Definition of User Experience**. Fremont, CA, USA. Nielsen Norman Group, 2016.

NORMAN, D.; J. NIELSEN. Fermont, CA, USA. **The Definition of User Experience**. from <https://www.nngroup.com/articles/definition-user-experience/>. Retrieved 15.11.2016.

NORMAN, D. A. **The design of everyday things**. New York, USA: Basic Books, 2002.

RAU, P. P. **Cross-Cultural Design. Cultural Differences in Everyday Life:** In: 5th International Conference, CCD 2013, Held as Part of HCI International, 2013, Las Vegas, NV, USA, July 21-26, 2013. v.5.

ROBIER, J. **Das einfache und emotionale Kauferlebnis: Mit Usability, User Experience und Customer Experience anspruchsvolle Kunden gewinnen**. Berlin: Springer, 2015.

ROGERS, W. A.; A. D. FISK. **Toward a psychological science of advanced technology design for older adults.** In: The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, Oxford, 2010.

ROUSE, M. Newton, MA, USA. **Definition Web Site**. from <http://searchmicroservices.techtarget.com/definition/Web-site>. Retrieved July 2016.

THE HENRY FUND. **Online Travel Agencies - Consumer Discretionary Sector**. Henry B. Tippie School of Management. Iowa City, IA, USA, 2016.

THE WORLD BANK. Washington D.C., USA. **Internet users (per 100 people).** from <http://data.worldbank.org/indicator/IT.NET.USER.P2?end=2015&start=1990&view=chart>. Retrieved 31.01.2017

THÜRING, M.; S. MAHLKE. **Usability, aesthetics and emotions in human–technology interaction.** In: International Journal of Psychology, Minneapolis, MN, USA, 2007. p. 253-264. v.42.

TOP10LIST. Top 10 **Uses of Internet. Feature Technology.** From <http://top-10-list.org/2013/06/22/top-10-uses-of-internet/?utm_source=feedburner&utm_medium=feed&utm_campaign=Feed%3A+top-10-list%2FIAGi+(Top+10+List)>. Retrieved 05.06.2016.

U.S. DEPARTMENT OF HEALTH & HUMAN SERVICE. Washington D.C., USA. **Usability Evaluation Basics.** from <http://www.usability.gov/gsearch/qualitative>. Retrieved 21.01.2016.

APPENDIX A – SUS QUESTIONNAIRE

Figure 45 – SUS Questionnaire (BROOKE, 1996)

Participant ID: _____ Site: _____ Date: ___/___/___

System Usability Scale

Instructions: For each of the following statements, mark one box that best describes your reactions to the website *today*.

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

Please provide any comments about this website:

APPENDIX B – SUPPOORTING SHEET FOR HOTEL BOOKING PROCESS

The following figure 46 illustrates the supporting sheet that was given to every user during the online booking process.

Figure 46 – Supporting sheet during hotel booking process

Book a hotel

- Where: **Rio de Janeiro, Brazil at Copacabana**
- When: **December 15th till December 17th 2017**
- What: **2 rooms with 4 persons in total**
- **Breakfast included**
- < **$1.000** in total
- **No credit card required** for booking
- **No Prepayment**
- **Free Cancellation**

Personal Information:
Christian Marian
15/12/1992
Telephone: (11)949604254
Email: christianmarian92@googlemail.com

→ Cancel immediately

Send text to customer service as a Feedback

Dear Booking Customer Support,

I apologize for the cancelling my reservation but this is a university study and I didn't aim to travel to Rio de Janeiro, Copacabana on these dates. Anyway, as soon as I travel to Rio, I will book on your website and I intend to travel to the hotel I just cancelled.

I am sorry for the inconvenience

APPENDIX C – AUTO EVALUATION QUESTIONNAIRE

This questionnaire is based on John Brooke's (1996) SUS questionnaire and on Bastien and Scapin's (1993) criteria.

Figure 47 – 20-question-survey adopted to Brooke (1996) & Bastien and Scapin (1993)

	Strongly disagree				Strongly agree	
	1	2	3	4	5	N/A
1 I liked using the website.						
2 I felt comfortable using the website.						
3 I thought this website was easy to use.						
4 I did not need any help to complete the task.						
5 I found this website very pleasureable to use.						
6 I would imagine that most people would learn to use this webiste very quickly						
7 I liked the functions offered by the website, such as filters.						
8 The website guided me nicely through the process.						
9 I always had control in the process. I never felt helpless or clueless.						
10 The website supported me in completing the task (e.g. calaneder popped up by date entrence)						
11 There was neither too much nor too little information on the screen.						
12 Text was comfortable and easy to read.						
13 I think the website was efficient. (No unnecessary steps or windows needed)						
14 I like the labels, symbols, text size and colors on the website.						
15 It was easy to identify the right buttom I was looking for.						
16 The website provided a good support when problems occurred.						
17 The error messages were very specific.						
18 The error message helped me to correct the error.						
19 The effort required was proportionate to the goal.						
20 The achieved result met my prior expectations.						

APPENDIX D – FREQUENTLY ANSWERED QUESTIONS

In the following table 31 the frequently answered questions of critical users are listed. Since the Likert scale, which ranges from 1 to 5, is subjective, the individual worst rating is the worst rating given in the whole questionnaire by the specific user. The absolute worst rating is 1 but if a user never gave this rating, and instead he rated some questions with 2, this will be considered as the worst rating. This research assumes that the worst rating by a critical user (even if it is not a 1) means the user is upset in some way which makes it possible to consider it as the worst rating. The same applies to the best rating.

The first column always shows the question that was rated as worst (or as the best) by at least one critical user. The second column shows how many critical users opted for the worst/best rating. And the last column shows which user rated this question.

Table 31 – Frequently Answered Questions

Frequently Answered Question					
Worst Ratings *(by critical user)*			**Best Rating** *(by critical user)*		
Question	Amount	User	Question	Amount	User
4	3	14, 16, 34	7	3	18, 19, 34
8	1	18	10	1	34
9	1	18	11	1	19
11	1	29	12	1	34
13	2	18, 29	13	2	16, 19
14	1	19	14	1	16
15	2	18, 19	17	1	14
16	2	29, 34	19	3	16, 19, 29
17	2	16, 34	20	2	19, 29
18	2	18, 34	-	-	-
19	1	18	-	-	-
20	2	18, 34	-	-	-