Oscar Ranta

Analyzing the accessibility and usability of the PSOP system through user testing with visually impaired users

Master's Thesis in Information Systems
Supervisors: Prof. Anna Sell, Åbo Akademi
Jani Grönberg, Kuntien Tiera Oy
Faculty of Social Sciences, Business and Economics
Åbo Akademi University

Åbo 2019

Master's Thesis in Information Systems	
Author: Oscar Ranta	
Title: Analyzing the accessibility and usability of the PSOP system through user testing with visually impaired users	
Supervisor: Anna Sell	Advisor: Jani Grönberg

Abstract:

The demand for an accessible Web continues to increase and over the last decade, efforts have been made to establish Web accessibility for everyone, regardless of disability. However, the underlying problem for an inaccessible Web, is that the content is not designed according to the mindset of people with disabilities. The inaccessible perception of the Web increases as barriers such as inadequate structure, lack of image description and misleading information are present.

In this master's thesis the interface of the PSOP system (Palveluseteli- ja ostopalvelujärjestelmä) was investigated. The study was conducted in collaboration with Kuntien Tiera Oy in the hope of testing the system with visually impaired people. The goal of the study was to identify both usability and accessibility related problems that visually impaired people face while interacting with the system. Additionally, the importance of the study arises due to the fact that the Finnish Government has determined Web accessibility guidelines for content provided by the public sector. Therefore, another goal was to discover current problems with the user interface and to analyze how well the system satisfies the *Accessibility Directives*. The demand of identifying the problems increases significantly, as the PSOP system is going through development after the second quarter of the year.

A user-centered test approach was established to achieve the goals. Eight visually impaired users participated in the tests. The tests were conducted according to the participants' requirements in order to establish a realistic interaction with the system.

The analysis identified both usability and accessibility related problems affecting real users. According to the findings, visually impaired people will experience significant obstacles when interacting with the system. As a result, Kuntien Tiera Oy will need to take the problems into consideration in order to satisfy the *Accessibility Directives* and to ensure accessibility for all its customers.

Key words: Usability, Accessibility, User testing, Disability, WCAG

Date: 27.05.2019	Åbo Akademi

Table of Contents

List of figures

List of tables

1. Introduction

For over a decade, the importance of usability testing has increased significantly and remains as a central part of the design process as technological breakthroughs continue to impact companies and organizations. The methods for usability testing are not new and have changed over time. Consequently, a variety of methods can be utilized, and it is important to analyze which method is the most suitable one for a specific product or project. Undoubtedly, stakeholders and interested parties have acknowledge the value of usability and an increasing amount of time and money is now invested in order to find advancements to satisfy the customers user experience. (Morris & Dillon, 1996)

The demand for an accessible Web continues to increase. According to Kapsi, Vlachogiannis, Darzentas and Spyroual (2009), over the last decade, serious efforts have been made in order to establish Web accessibility for people with disabilities. However, a core issue for an inaccessible Web is the fact that the content is not designed according to the mindset of people with special needs (Bradbard & Peters, 2010). According to Bardbard and Peters (2010), the inadequacy of Web accessibility results from an illogical design process instead of from insufficient utilization of technologies. Consequently, barriers such as misleading content, inadequate structure, lack of image description and access to tables and other content are commonly encountered (Bradbard & Peters, 2010).

Although, improvements are constantly done to ensure Web accessibility, people with disabilities will still encounter difficulties when interacting with the Web. The importance of establishing Web accessibility will continue to increase, since new technology emerges and as users will keep interacting with web content in new and innovative ways (Kapsi, Vlachogiannis, Darzentas, & Spyroual, 2009). As a result, several countries are demonstrating their support towards Web accessibility, by policies and legislations that require authorities to make digital services accessible. Similarly, Finland has followed the directive of the European Parliament, to ensure

accessibility of websites and mobile applications provided by the public sector, and included the *Accessibility Directive* into the law, which entered into force on 1st of April 2019 (Valtiovarainministeriö, 2018). However, the *Accessibility Directive* concerning websites provided by the public sector, will be applied gradually from the 23rd of September 2019 (Valtiovarainministeriö, 2018).

Kuntien Tiera Oy has acknowledged that the *Accessibility Directive* concerns the company and the service voucher and purchased service system (Palveluseteli- ja ostopalvelujärjestelmä, PSOP). Therefore, Kuntien Tiera Oy is determined to address both usability and accessibility related problems, in order to comply with the law. Consequently, Kuntien Tiera Oy indicates that the customer interface is of high priority, due to the EU Accessibility Directive. Kuntien Tiera desires that the PSOP system would be tested with visually impaired users on the customer side, as further research is needed to identify problems and hindrances.

1.1 Context

Accessible Web design, universal usability and *design for all,* involves the design of products and services, as well as websites, to be usable by the general public or by anyone who wishes to interact with information on the Internet (Brophy & Craven, 2007). This involves the whole population in the society, not excluding the minority of people with disabilities. According to Brophy and Craven (2007), the requirements are satisfied when the on-screen information is accessible by all, regardless whether the user uses assistive technology or not. The founding father of the World Wide Web, Tim Berners-Lee, pointed out in 2001 that “the power of the Web is in its universality. Accessibility by everyone regardless of disability is an essential aspect” (Sevilla, Herrera, Martínez, & Alcantud, 2007). Searching for information on the Web, or simply accessing it without any significant hindrances, might be something that we take for granted. However, the ageing population and people with different kind of disabilities, e.g. sight impairment, may experience real obstacles. A considerable amount of time and work has been done in order to make

the Web more accessible for the purpose of utilizing the Web to its fullest, whether the user has a certain disability or not.

A substantial number of guidelines have been set by the World Wide Web Consortium in order to ensure accessibility on the Web. These guidelines are referred to as Web Content Accessibility Guidelines (WCAG). The initial version, WCAG 1.0, was introduced in 1999 when the Web was not as heavily used as in the 21st century and mostly consisted of informative sites and a handful of e-commerce platforms. Today, the accessibility guidelines have been further developed, due to the digitalization of the society and a new set of guidelines, including the usage of mobile devices, known as WCAG 2.1, has been introduced. The Finnish government has consented to making the digital services provided by the public sector accessible for all the citizens, in order to benefit the whole population. A proposal from the government of a law that ensures an accessible Web has already been submitted to the Parliament, with the intention to promote equal opportunities to act fully in a digital society. The accessibility requirements of public websites and services begin to apply gradually from the 23rd of September 2019, depending on when the website is published. These accessibility guidelines concern Kuntien Tiera Oy considerably and are highly relevant for the PSOP system. Since the PSOP system provides an effective alternative to utilize and manage social welfare and health care services on their platform, an accessible platform is required. Kuntien Tiera Oy is determined to increase as well as test for the accessibility of their platform, due to the fact that the customer base consists of an ageing population and of people with disabilities using a variety of assistive technology. Nonetheless, a thorough, more technical study regarding the Web Content Accessibility Guidelines has initially been conducted by Kuntien Tiera Oy. However, it is important to point out that a fully accessible platform that satisfies the accessibility guidelines does not necessarily reflect whether the platform is actually accessible by the end user. Therefore, Kuntien Tiera Oy hopes to test if the platform is accessible in reality, by conducting a usability test involving real users with disabilities.

1.2 Purpose of the study

The purpose of this study is to understand what challenges the users face when interacting with the PSOP system. Furthermore, it is important to analyze what kind of hindrances the target user group face and whether the system holds obstacles and problems that are related to accessibility and usability.

Within the scope of the assignment from Kuntien Tiera Oy, it is hoped that the PSOP system will be tested for its accessibility with real users. The usability testing regarding the customer's accessibility has the highest priority and is therefore the center of the study. In the customer interface, high priority is given to accessibility and the EU Accessibility Directive. Earlier feedback on the system has focused on problems among the visually impaired users on the customer side, and further research is desirable to address these problems. Since a large amount of money is invested yearly in the service, the thesis will work as an indicator and as a "checklist", in order to recognize the systems current state and for Kuntien Tiera Oy to determine what actions are required for establishing accessibility for all.

The study is conducted both through a survey and a usability study to analyze if the PSOP system satisfies the WCAG guidelines that are defined for web services provided by the public sector. The corner stone of the study is to identify how accessible the system is among disabled users, even if WCAG is met or not, since it does not necessarily mean that a service or platform is accessible even though all the guidelines would have been satisfied.

1.3 Identifying the research questions

Usability and foremost the accessibility of websites provided by the public sector are of significant importance, in order to ensure equivalent service for the whole society. Hence, the Finnish government has established guidelines for the gradual implementation of the Web Content Accessibility Guidelines that affect websites

provided by the public sector (Valtiovarainministeriö, 2018). The motivation behind the study, is to provide valuable insights for Kuntien Tiera Oy, in order to ensure and establish equivalent accessibility for all its users. The research will be conducted by user testing, involving participants with sight impairment who are utilizing assistive technology. In order to identify the problems within the system, the following research questions will be answered:

RQ1: *Although a technical test regarding the WCAG has previously been conducted, what kind of problems do users with sight impairment experience when interacting with the PSOP system?*

RQ2: *What are the reasons behind the identified problems within the PSOP system and what kind of actions should be done in order to eliminate the problems?*

2. Literature review

The following chapter will present the literature used in this study. Chapter 2.1 will discuss disability and how it is defined worldwide as well as in Finland. Chapter 2.2 will present the fundamentals behind Web accessibility. Chapter 2.3 will present the key aspects of Web Content Accessibility guidelines, how they have changed over time and how they have evolved into a standard that has been adopted and applied in the Finnish law. Chapter 2.4 presents the background behind usability and highlights the importance of it. Chapter 2.5 further discusses commonly used usability methods and characterizes differences between human centered and automated methods.

2.1 Defining disability

Defining disability under one general term, can be somewhat challenging. According to researcher's, disability is established and conceived when our society considers that the vast population is fully functional. On the contrary, we often fail to account the diversity among the population in our society. Since the definition of disability is a broad term and due to the fact that disability comes in many shapes and forms, it is difficult to characterize it under one specific term. In many cases, disability can be viewed as a part of the human condition, a condition that will most probably affect the majority of the world's population at a given point in their lives. This probability will seemingly increase at a later stage in life. On the other hand, the World Health Organization defines disability as an umbrella term that consist of various impairments, activity limitations and participation restrictions. According to the World Health Organization, the impairments in relation to health, are considered to be a flaw or loss of psychological, physiological, biological structure or body function. In its simplicity disability refers to a problem in an individual's body function. However, activity limitation is known as a limitation or difficulty that an individual is facing when executing a certain action, whereas participation

restriction is known as problems encountered in life situations. (World Health Organization & World Bank, 2011)

The number of people in our society with disabilities is increasing. This may in fact be due to the ageing of the population and a result of growing numbers of chronic health conditions that are related with disabilities. The chronic health conditions associated with disabilities are for example diabetes and mental illness to name a few. According to the World Health Organizations report on disability, there are more than a billion people in the world living with some kind of disability, which equals to roughly 15 percent of the world's population. These numbers are based on the 2010 global population estimate, with a growth of nearly 5 percent compared to the earlier WHO estimate from the 1970s. The condition of disability varies significantly and although it is estimated that approximately 15 percent of the world's population are living with some kind of disability, the World health Organization reports that there is no explicit or implicit distinction between health conditions when it comes to classifying disability. However, not all of these 15 percent of the world's population, are living with something that is classified as severe disability, such as autism, dementia or Down syndrome just to name a few. Disability commonly correlates with disadvantage, as disability is often emphasized with wheelchairs and with blind and deaf people. However, the nature of disability varies significantly depending on different conditions and factors. (World Health Organization & World Bank, 2011).

According to the Finlex data bank of the Act on Disability Services and Assistance (3.4.1987/380 §2), a disabled person is referred to a person who, due to an injury or illness, has a particular difficulty in normal life performances and activities over a longer period of time. In other words, disability is taking form when a disabled person faces distinct hindrances, which restricts their effective presence in society equally compared with others.
The late theoretical physicist and cosmologist, Professor Stephen W Hawking (World Health Organization & World Bank, 2011), speaks about disability and equal enjoyment of human rights in the society as follows:

"Disability need not be an obstacle to success. I have had motor neuron disease for practically all my adult life. Yet it has not prevented me from having a prominent career in astrophysics and a happy family life."

(World Health Organization & World Bank, 2011)

According to Henrik Gustafsson from Invaliidiliitto (2016), disability is no longer defined primarily as a medical or social issue, but above all as a human rights issue. Thus, human rights are seen as the basis for new legislation and for people with disabilities as human rights holders, who must also be able to claim their rights as individuals. The United Nations Convention on the Rights of Persons with Disabilities (CRPD), adopted by the United Nations General Assembly in late December 2006, is the first comprehensive human rights treaty of the 21st century (United Nations, 2006). Finland was among the first countries to sign the agreement in early 2007, as the Parliament approved the government's proposal for ratification (Gustafsson, 2016). The CRPD came into force on the 10th of June 2016 (Gustafsson, 2016). The agreement quickly became one of the most widely accepted UN human rights conventions and has already been signed by over 160 parties, including the European Union as a regional integration (Mahlamäki, 2015). The key distinction in the convention is that disability does not reside in the person, rather in the society. The United Nations Convention on the Rights of Persons with Disabilities will be discussed more in depth in a later stage.

In order to draw a broader perspective on the statistics of disability in the European Union, data retrieved from the European Health and Social Integration Survey (EHSIS), illustrates the working ability among disabled people, their well-being and their participation in restrictive barriers (Eurostat, 2015b). The presence of limitations and the well-being of people with disabilities, compared to other people, can be examined from the indicators of the survey (Eurostat, 2015b). In Finland, the corresponding number of people with disabilities between ages 15 and 64, has risen to 15 %, which equals to over 500 000 of the whole population (Eurostat, 2015b).

According to Eurostat (2015b), the responding percentage of the sampled population was approximately 26 %.

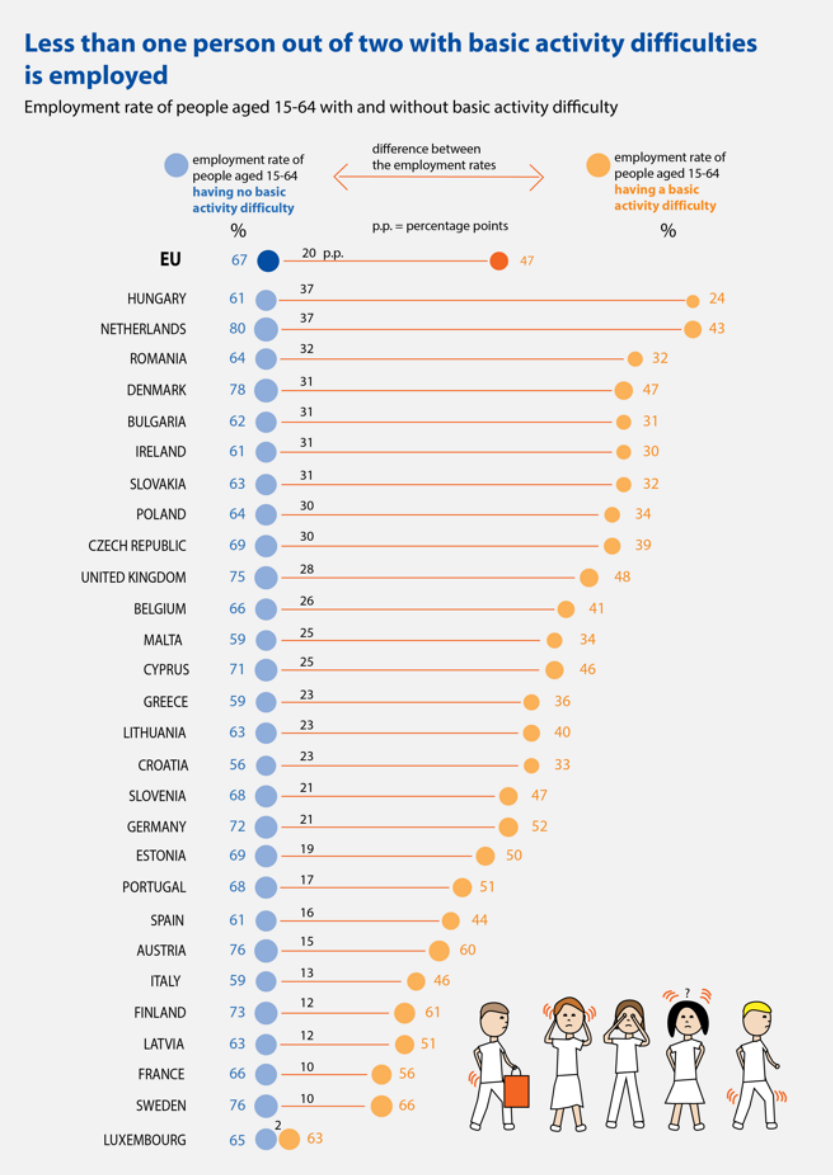
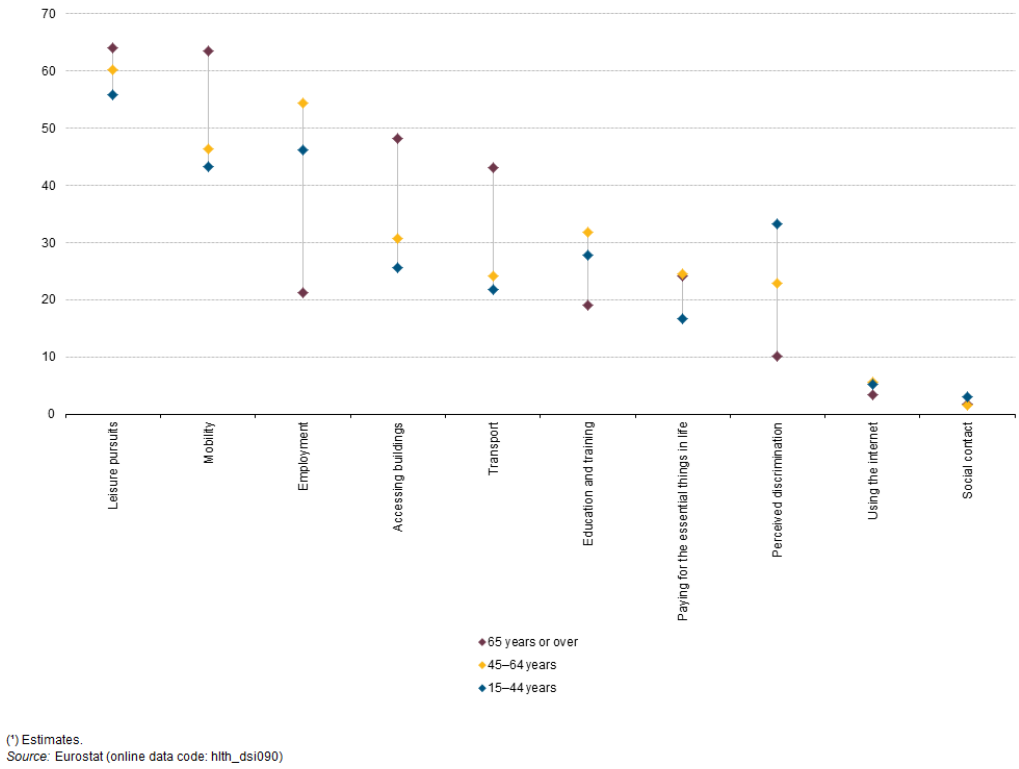


Figure 1: Employment among people between ages 15-64 with activity difficulties in European countries (Eurostat, 2015b)

Figure 1 illustrates that less than every second person with a basic activity difficulty, between ages 15 and 64, are employed. Such activity difficulties are for example hearing, seeing or a similar difficulty that has lasted or is expected to last more than six months. According to figure 1, the percentage of employment among people

with no activity difficulties in Finland is 73 %, which is among the highest in the European Union. However, the employment rate of people with a basic activity difficulty in Finland is an astonishing 61 %, barely behind our neighboring country Sweden, with a rate of 66 %. The employment rate among people with a basic activity difficulty in Finland, is significantly higher compared to central European countries. The employment rate in Finland is also significantly higher than the average percentage in the European Union. Figure 1 further illustrates that Finland with 12 %, is among the countries with the lowest percentage difference between people aged 16 and 64 with and without a basic activity difficulty. (Eurostat, 2015b)

Figure 2 illustrates different difficulties in various life areas among disabled people. A specific life area that is significantly interesting for this research, is the use of Internet. According to figure 2, it can be noted that the use of Internet and age does not have a distinct correlation. Nonetheless, the use of Internet does not necessarily make a significant difference on the issue of using the Internet among disabled people. The relation between disabled people and their usage of different Internet services, works as a prelude for the upcoming discussion about Web accessibility. (Eurostat, 2015a)



(*) Estimates.
Source: Eurostat (online data code: hlth_dsi090)

Figure 2: Disabled persons in the European Union aged 16 and over, reporting hindrances in various life areas. (Eurostat, 2015a)

2.2 Defining Web Accessibility

The general term of accessibility addresses the equality of user experience for all people in the society, regardless of disability or age-related impairments. Accessibility is one of the key points in the CRPD, as the aim is to help and overcome limitations and hindrances that people with disabilities face when participating in various life areas. The corner stone of accessibility is to take the diversity of the people into account, as well as the design and the implementation of the environment. This includes the accessibility of different services, the usability of certain tools and foremost the comprehensibility of the information as well as the possibility to participate in personal decision making. Accessible environments are commonly implemented in the Finnish society, as most buildings and public transport have an easy access for people with various health impairments. (United Nations, 2006)

This study will, however, have its focus on Web accessibility. According to the Web Accessibility Initiative (2005), Web accessibility is referred to a Web that is accessible by anyone, not excluding various impairments or age restrictions. Web accessibility focuses on platform design, allowing people with disabilities to perceive, interact and use the Web as it is or with tools such as screen readers and other assistive technologies (WAI, 2005). As the founder of the World Wide Web and the director for the Web Accessibility Initiative, points out: *"the power of the Web is in its universality. Access by everyone regardless of disability is an essential aspect"* (Sevilla et al., 2007). In other words, the essential condition of the Web should be designed to work and practiced by everyone, regardless of disabilities such as mobility, hearing, visual or cognitive impairments (WAI, 2005). Although the usage of the Web further increases communication as well as interaction, the Web Accessibility Initiative (2005), points out that websites and digital services, which are poorly designed, create usage barriers that may exclude the ageing population and people with disabilities from using the Web. An earlier study conducted by Nielsen Norman Group, revealed that a person with visual impairment, low vision or completely blind, experiences the Web three times as hard to use, compared to sighted users (Brophy & Craven, 2007). Consequently, meaning that visually impaired people have a significantly higher possibility to be excluded from accessing the Web. In order to ensure that people with disabilities are able to digitally participate in various life areas (i.eg. education, employment, services and health care), is according to the United Nations Convention on the Rights of Persons with Disabilities, foremost a basic human right (WAI, 2005). In other words, accessibility has been relocated from a highly immense theoretical aspect into a significantly important goal, embraced by several European countries, including Finland.

The importance of accessibility has increased significantly in recent years, due to the substantial development of digitalization (United Nations, 2006). Today, accessibility is perceived as a political issue, instead of an issue related to web design, and further registered in the CRPD (United Nations, 2006). According to article 9 of the CRPD (United Nations, 2006), the issue regarding Web accessibility is defined as follows:

"States Parties shall take appropriate measures to ensure to persons with disabilities access, on an equal basis with others, to the physical environment, to transportation, to information and communications, including information and communications technologies and systems, and to other facilities and services open or provided to the public, both in urban and in rural areas"

(United Nations, 2006)

"Accessible Web design", *"design for all"* and *"universal design",* are all terms closely related to Web accessibility. These terms are compiled when the on-screen information is accessible and provided in a manner that does not contend with the use of assistive technology. According to Brophy and Craven, in terms of Web accessibility, *"design for all"* is specified as a website or web service that is accessible to everyone, and where a well-designed website does not have to scarifice its savvy features for accessibility. Undoubtedly, accessible web design benefits the general public, the importance of accessibility is increasing due to people's usage of the Web with multiple devices, such as personal computers, smart phones and tablets. However, social aspects are not the only factors that motivates the Web to be more accessible, financial factors are also a major key in the entirety to establish an accessible Web presence. According to Brophy and Craven, organizations and companies are realizing the growing demand of accessible websites and are therefore determined to promote their services to a wider audience. Taking this into account from an economic point of view, a Web that is accessible and further supports assistive technology, is a valid aspect to for a company to consider, in order to maintain and enhance the image of a business. Several organizations have already established this and are readjusting their strategies and policies in order to provide equal service for all. (Brophy & Craven, 2007)

Although an accessible Web benefits the general public, financial disadvantages are still present, as Richards and Hanson (2004) points out, *"the cost of creating accessible Web content is substantial".* Hence, as it may result in unexpected expenses for several websites that contain a substantial amount of aged content

that needs redesign, in order to be accessible. Additionally, an inconsequential improvement or effort in designing a website more accessible, e.g. labeling *ALT* tags for images, may promptly result in high costs for a company (Richards & Hanson, May 17, 2004).

2.3 Web Content Accessibility Guidelines

The goal of the World Wide Web Consortium's Web Accessibility Initiative (WAI), was in the late 1990s to improve the accessibility of the Web for people suffering from different disabilities. The vision was to provide an international standardized model for defining Web content accessibility. These guidelines are also known as the Web Content Accessibility Guidelines (WCAG) and where the first version, WCAG 1.0, was published in 1999 when the Web consisted mostly of informative sites and a handful of e-commerce platforms. The guidelines of WCAG 1.0 were mostly set in order to make HTML pages more accessible for disabled people, due to the rising development and need of assistive technology. (Reid & Snow-Weaver, Apr 21, 2008)

2.3.1 WCAG 1.0

According to the World Wide Web Consortium (W3C) (1999), WCAG 1.0 is comprised of 14 guidelines and within the guidelines there are a total of 65 checkpoints. These checkpoints are mostly intended for web developers to describe how web content should be designed in order to establish accessible (W3C, 1999). Consequently, to ensure the accessibility of the Web, WCAG 1.0 also consist of priority levels, ranging from 1 to 3 (W3C, 1999). These priority levels are created to enlighten the importance of the checkpoints, as each checkpoint is assigned with a priority level in order to ensure that the content is as accessible as possible for the general public (Power, Freire, Petrie, & Swallow, May 5, 2012). As a practical example, if the web content satisfies the requirement for all priority 1 checkpoints,

Level A has been established (W3C, 1999). Similarly, the updated version of the Web Content Accessibility Guidelines, WCAG 2.0 (which will be covered in a later stage), Level A is of significant importance for a websites accessibility and should always be required (Reid & Snow-Weaver, Apr 21, 2008). Furthermore, if all priority 1 and priority 2 checkpoints are fulfilled, the websites conformance level results in AA (W3C, 1999). Additionally, conformance Level AAA is established when all priority checkpoints are met (W3C, 1999). Noticeably, WCAG 1.0 was quickly recognized by several countries and for years it served as the standard for Web accessibility, as initially intended by WAI (Power et al., May 5, 2012). According to Power et al. (2012), despite WCAG 1.0 was widely accepted as the standard for Web accessibility, certain criticism arose against WCAG 1.0, resulting in loss of support for the guidelines. Most of the criticism arose opposing the testability of the web contents accessibility and towards the inconsistency of the technical knowledge of the guidelines (Power et al., May 5, 2012). According to Power et al (2012), the conformance levels missed serious empirical evidence, as difficulties arose in determining whether a website with a level AAA of conformance was in fact more accessible than a website with a level A. However, the awareness of the soon to be outdated guidelines of WCAG 1.0, served as an initiative for W3C in developing new guidelines for Web accessibility, WCAG 2.0 (Reid & Snow-Weaver, Apr 21, 2008). These guidelines were later published in 2008 by WAI, with the aim to resolve and address problems that the former guidelines contained.

2.3.2 WCAG 2.0

WCAG 2.0, which is also an ISO International Standard, ISO/IEC 40500:2012, was published in December 2008 (W3C, 2005). Since its publication, WCAG 2.0 has been acquired by several governments and organizations worldwide (W3C, 2005). WCAG 2.0 has rapidly been established as the standard for Web accessibility (Council of the European Union, 2016). In late 2016, the European parliament approved the directive and the guidelines of WCAG 2.0, which requires websites operated by the

public sector, to conform at least with the level AA of WCAG 2.0 (Council of the European Union, 2016).

The two main characteristics that distinguishes WCAG 2.0 from the prior outdated guidelines, are testability and further the non-reliability or dependence of a certain technology (W3C, 2008). In terms of testability, compared to the prior WCAG 1.0, which consisted of more technical guidelines, the main criteria for WCAG 2.0 is that the web content should not only be technically testable, but also testable by real physical users (Reid & Snow-Weaver, Apr 21, 2008). According to Reid and Snow-Weaver (2008), it is necessary to involve the human factor as well, especially when evaluating if the web content is competent enough to pass the criteria of the guidelines. However, WCAG 2.0 has a slightly different structure compared to the prior WCAG 1.0 and according to W3C (2008), WCAG 2.0 consists of guidance's, principles, specific guidelines and of success criteria. Additionally, W3C (2008) provides four specific principles for Web accessibility; *Perceivable, Operable, Understandable* and *Robust*, which can be seen in figure 3:

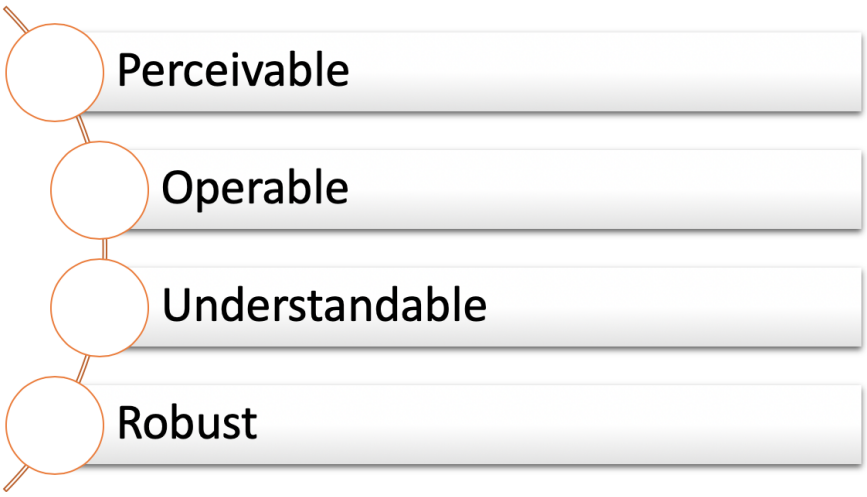


Figure 3: Principles of WCAG 2.0

According to W3C (2008), the principle of **Perceivable**, can be summarized as the aspects of Web accessibility, which requires the content and the information to be perceivable by all the senses of the user, free from undetectable content (W3C,

2008). In practice, it means that the web content must be redesigned and capable of being converted into braille, symbols, speech or to a simplified language (W3C, 2008). Consequently, the web content should be accessible, regardless of the underlying disability. According to W3C (2008), the principle of perceivable has a total four guidelines and are summarized as follows:

- 1.1 Provide text options for web content that is not in text format and further be transformed into content that certain people require, i.e. audio and uncomplicated language, just to name a few.
- 1.2 Provide options for media content as a function of time
- 1.3 Provide web content in varying ways, not excluding any information or without losing the structure of the content
- 1.4 Design the web content in a specific way that allows for efficient interaction, excluding hindrances with the content.

According to W3C (2008), the principle of **Operable***,* signifies that the web content needs to be programmed and designed in an efficient way, in order for people with disabilities to use the Web to its full extent (W3C, 2008). The user interface, its components and the navigation on the Web must be operable, signifying that all functions on the site must be accessible by only using the computers keyboard (W3C, 2008). Additionally, the content must be as clear as possible, since headers must be descriptive and contain a clear link that informs the user if it can be interacted with W3C (2008). W3C (2008) further points out, that there should not be time constraints nor anything on the website known to cause seizures, such as epilepsy. W3C (2008) summarizes the four guidelines of operable as follows:

- 2.1 Provide full functionality and accessibility to content by the usage of keyboard
- 2.2 Provide web content with no time constraint and ensuring users to sufficiently read and use the content
- 2.3 Provide web content that does not cause seizures

- 2.4 Providing ease of navigation throughout the content and ensuring that users are able find the desired information as well as locate themselves on the site

The content has to be **Understandable** and according to W3C (2008), WCAG 2.0 clearly indicates that the content on the platform must be programmed in such a way that people with disabilities are able to understand it (W3C, 2008). In other words, information, various functions and the language must be clear and understandable (W3C, 2008). Additionally, if the web content includes forms, the forms should clearly indicate the formality in which way the information is desired (W3C, 2008). In accordance with W3C (2008), the principle of understandable only contain three guidelines, which are summarized as follows:

- 3.1 Provide web content that is readable as well as understandable
- 3.2 Provide predictable web content, in order to ensure that users can predict what is going on
- 3.3 Provide input assistance, in order for users to avoid mistakes and possible correct them

According to W3C (2008), the web content has to be **Robust** and programmed in order so that users with assistive technology are able to access the information with their tools (W3C, 2008). The principle of robust is significantly important, as the technology behind assistive tools and devices is constantly evolving. The principle only consists of one guideline and according to W3C (2008), it can be summarized as follows:

- 4.1 Maximizing compatibility to current agents as well as the ones to come, by highlighting the importance of assistive technology and tools.

The fundamental idea behind the twelve guidelines of WCAG 2.0, is to provide web-developers and designers an underlying path that needs to be followed in order to produce accessible web content. Noticeably, WCAG 2.0 also consist of conformance levels, similarly ranging from Level A to Level AAA. Consequently, Level A should

always be required. However, it is significantly important that all the success criteria are testable, as difficulties may rise when determining if the web content meets the success criteria or not. Consequently, if the success criterion is satisfied, it may in fact indicate that a website does not support assistive technology, which is an essential aspect when discussing Web accessibility. According to W3C, the term testable, in terms of WCAG 2.0, consists both of machine testing and human interaction in order to evaluate the web content. However, the human factor and the importance of performing user tests with real users increases, as a conformance Level of AAA does not necessarily indicate that the web content is fully accessible to all individuals. (W3C, 2008)

In addition, WCAG 2.0 consists of particular techniques; sufficient- and advisory techniques, as well as of failure related techniques. Sufficient techniques consist of reliable ways to achieve the success criteria, whereas advisory techniques are commonly suggestive actions that are needed in order to improve accessibility. Additionally, failure related techniques, are listed as causes for not satisfying the success criteria of WCAG 2.0 and are documented in order to avoid them in the future. W3C further suggests, that researchers should consider applying all layers of guidance; *principles, guidelines, success criteria* as well as the different techniques, in order to improve the possibility of providing web content that is more accessible for the general public. (W3C, 2008)

Consequently, as testability is one of the corner stones of WCAG 2.0, usability is strongly present and an aspect that needs to be accounted for. Undoubtedly, WCAG 2.0 is usability-oriented with specific research-based guidelines (Termens et al., Apr 20, 2009).

According to Valtiovarainministeriö, Finland strictly follows the Accessibility Directive of the European standard EN 301 549 (Accessibility requirements suitable for public procurement of ICT products and services in Europe), regarding the level of conformances and the accessibility of the websites provided by the public sector. Consequently, all web services provided by the public sector in Finland, are demanded to satisfy at least conformance Level AA. (Valtiovarainministeriö, 2018)

2.3.3 WCAG 2.1 and applying the Accessibility Directive in Finland

The new version of the Web Content Accessibility Guidelines contains no drastic changes compared to the previous version, which has been the first and most central web guide since 2011. However, WCAG 2.1 extends the former guidelines with 17 new criteria, of which 12 are required for Level AA. Undoubtedly, as technology is constantly evolving, WCAG 2.1 also contains guidelines for mobile devices and applications and is expected to contribute cognitive and mobile accessibility. The new set of guidelines also include instructions for contrasting visual elements, line-up between functions and for highlighting features on the web content with assistive technology. In addition, WCAG 2.1 is fully compatible with WCAG 2.0, as it follows the same structure, principles and conformance levels as the former guideline. As of 2019, the initial guidelines of WCAG 1.0 will be replaced and overwritten by the new WCAG 2.1. (W3C, 2018)

Although WCAG are considered to be a valuable set of tools for developing and evaluating web content, limitations and challenges are still present. The guidelines cover mainly the technical availability of the content and do not necessarily account for usability. The aspect of involving usability, has been under development ever since the updated version of WCAG was published. However, improvements have been made in order to include usability as a key point in WCAG. Studies show that designing a website according to WCAG, solves occasionally half of the problems that visually impaired users face. Consequently, WCAG 2.1 has noted this aspect and improvements in previous ignorance towards learning difficulties as well as cognitive and linguistic limitations and other disabilities have been made. (W3C, 2018)

The Accessibility Directive and the resulting national legislation, require authorities in Finland to design their digital services available to all citizens, not excluding people with disabilities. The proposal from the Finnish government, concerning a

law on the provision of digital services, has already been accepted and entered into force on 1st of April 2019. The accessibility requirements for public websites and services begin to apply gradually from the 23rd of September 2019. The purpose of the accepted Accessibility Directive is to provide accessible digital services that benefits the whole population. The government of Finland intends to promote equal opportunities to act fully in a digital society and thereby follow the minimum requirements established by the European Union, concerning the accessibility of public administrative websites and mobile applications. This is an important aspect to consider, as the law demands the public sector to provide its customers with an equal opportunity to access authorities digitally. According to Valtiovarainministeriö, the Accessibility Directive applies for content on websites and mobile applications provided by the public sector, as well as for organizations that manages public administration tasks. These are for example government agencies and businesses, municipal authorities, schools, universities as well as third party organizations that receive monetary assistance from the authorities to produce and maintain websites. Similarly, to the criteria of WCAG 2.1, the Accessibility Directive is divided into three requirement levels, *A, AA* and *AAA*. As a result, websites and web content provided by the public sector, must satisfy at least Level AA, in order to satisfy the Accessibility Directive. (Valtiovarainministeriö, 2018)

The gradual implementation of the Accessibility Directive in Finland is illustrated in Figure 4.

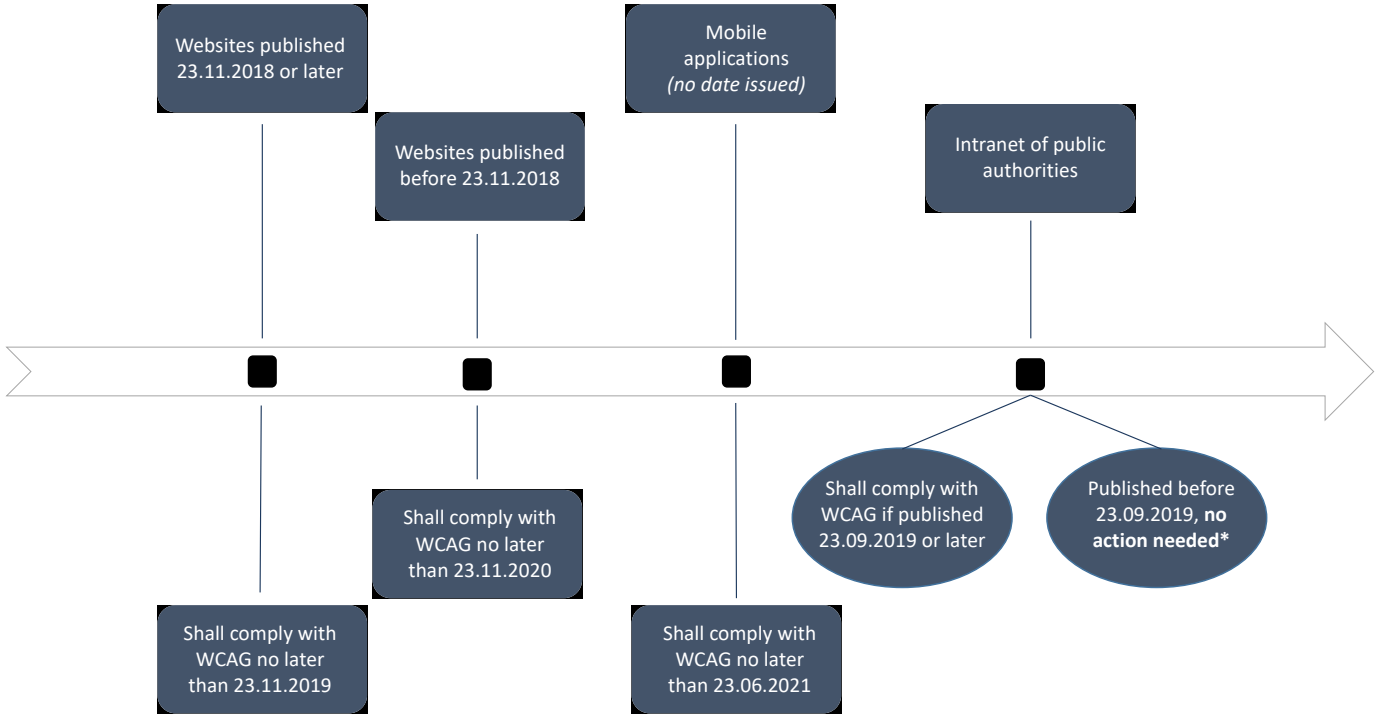


Figure 4: Gradual implementation of the Accessibility Directive in Finland

Figure 4 illustrates how the Accessibility Directive is gradually implemented in Finland. Websites published on September 23, 2018 or later must comply with the directive by September 23, 2019 at the latest. Websites published before the September 23, 2018 must comply with the directive by September 23, 2020 at the latest. However, mobile applications and mobile content that is provided by the public sector must meet the requirements by June 23, 2021. Additionally, Intranets of public authorities and public law establishments, including Intranets used in workplaces, must meet the requirements if they are published on September 23, 2019 or later. However, for Intranets published before September 23, 2019, accessibility requirements apply only if the entire Intranet is renewed or upgraded.

2.4 Defining usability

Usability is a central part of this study and closely related to accessibility as well as with the combined terminology universal design. The essential aspects regarding usability, will be thoroughly discussed in the following chapter. The chapter will discuss what usability is, how it is defined, how to evaluate it and how it is taken into account within the framework of this study.

Usability has been a highly utilized evaluation tool for a long period of time, ranging back to the initial interaction between users and systems. Usability has been a corner stone of human-computer interaction (HCI) and is defined as *how well a product or system can be used by a specific user*. Terms such as human-computer interaction, user interface and user experience are all closely related to usability, however, there are some noticeable differences. Human-computer interaction can be translated as the interaction between the user and the computer or when man and computer meet. The interaction consists of both users and computers, providing outputs and receiving inputs (Dix, Finlay, Abowd, & Beale, 2014). Though usability and HCI are closely related, Dix et. al. (2014) argue that an effective discipline of HCI interaction should not be strictly based on usability analysis and that the evaluation methods vary. Evaluation methods for usability lie with its own niche which is included in the larger framework (Dix et al., 2014). In comparison, one can relate to the fact that HCI and user interface are closely related and are characterized by similar values. User experience, however, is more closely related to usability. According to Nielsen (2016) usability is concretely a quality attribute of user experience (UI) and mostly dependent on the user's preferences. One has to note that though the user interface might fulfill all the requirements set by usability components, the user experience might not be satisfied if it misses functions or utility.

The definition behind usability has varied over time and can be viewed from several perspectives. Generally, usability is a concept that can be implemented in more or less anything that we use. However, in this thesis we will lay our focus on web

usability and the view of usability among systems. One can easily note that something that is pleasant and easy to use is more usable and owns a good usability, compared to products that tend to annoy the user. According to Shackel (1990), another perspective behind usability is "*That systems should be easy to use, easy to learn, flexible, and should engender a good attitude in people"* (Benyon, 2014). Alternatively, usability similarly refers to the quality of the interaction, where parameters such as duration of performed tasks, number of errors occurred as well as measuring how long it will take for a user to become a competent or qualified user. Nielsen (2010), however, points out the importance of realizing that there is not a specific nor one-dimensional property of a user interface, when referring to usability. In all its simplicity, one can argue that usability is the measurement of quality that a user experiences when he or she is interacting with a specific system.

The study will adopt to the ISO 9241-11 definition of usability, which is the international standard for usability and also known as the *Ergonomics of human-system interaction.* According to the International Organization for Standardization, usability is defined as the "*extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use"*. (ISO 9241-11, 2018)

Furthermore, to obtain a wider perspective concerning usability, Nielsen (2012) points out that usability is furthermost a quality attribute as well as a guideline of methods to improve the usage within a process. A quality attribute can be defined as an attribute that evaluates how easy or difficult a certain user interface is actually to use and interact with (Nielsen, 2012). According to Nielsen (2012), usability has multiple components and can simply be defined by five quality components which are illustrated below in figure 5.

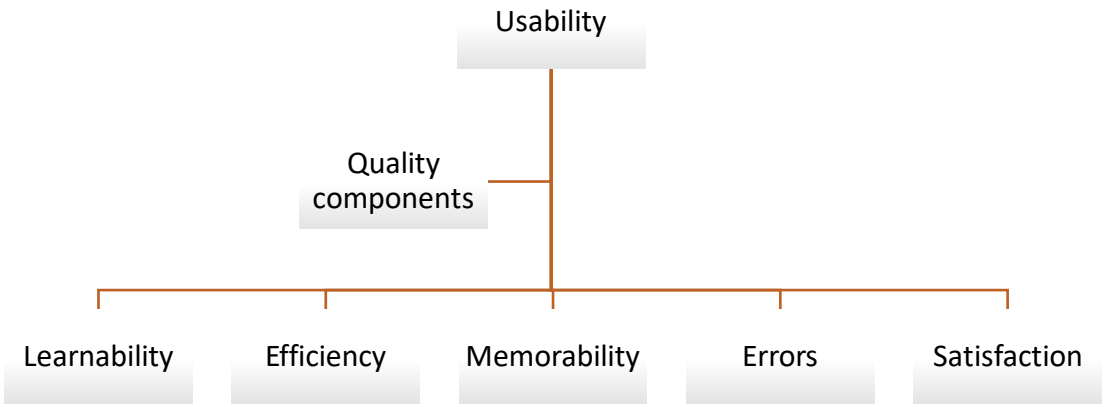


Figure 5: Five quality components of usability

Figure 5 illustrates five unique quality components that define usability. **Learnability** is user oriented and measures how easy or difficult it is for a first-time user to complete a given task. Furthermore, learnability characterizes the ease and effectiveness for a user to learn and operate with the system. The quality components indicate that a system needs to be designed in a specific way, so that the user in question can learn to use the system without obstacles. **Efficiency** is referred to user capability or to the amount of resources consumed when operating a system. Efficiency can be viewed as a second step after learnability, where it measures how efficient the user is to perform a given task, after he or she learned to use the system. As Nielsen (2012) notes: *"Once users have learned the design, how quickly can they perform tasks?".* A high level of productivity with the system is preferable, when the user managed to cope with the system. A system should be ***memorable*.** In other words, **Memorability** is reached when the user can return without obstacles to interact with the system, even though he or she would have been absent from using the system for a certain period of time. According to the quality component, the user should be able to reestablish proficiency and not be forced to learn the system all over again. The components further characterize **Errors**. Errors are difficult to completely exclude and sometimes inevitable. The

quality components point out that a system would need to have a low rate of errors, in order for the user to be able to interact efficiently. The severity of errors should also be minimal for users to be able to recover and overcome them without major problems. The final quality component is **Satisfaction**. Satisfaction characterizes how enjoyable the interaction is with the system. The quality components are highlighting the importance of the system satisfaction, due to the fact that whether the systems do not satisfy the user's expectations or needs, he or she might lose interest and decide to quit. (Nielsen, 2012)

One can relate to the fact that usability rules the Web and the importance of a usable Web is inevitable. The similarity can be seen in everyday life situations, if a customer does not like a certain product or she cannot find it, she will not buy it. The same applies for web usability, if the quality components do not satisfy the user's needs, the possibility increases for the user to look for other alternatives or simply give up. A poor usability that does not satisfy the user's needs, increases the risk for the user to interrupt and leave. According to Nielsen (2012), users may leave for several reasons: if the system is difficult to use, if the user gets lost or if the content is difficult to understand are just a few reasons why users might leave. This aspect will become even more severe when discussing usability in relation with Web accessibility among disabled people. However, a substantial amount of research has been conducted about usability and, according to researchers, it is problematic to find a universal indicator or techniques for usability, as well as for its evaluation, that would be possible to apply to the vast majority of systems in need (Whiteside & Holtzblatt, 1988). However, due to the extensive usage and emerged number of systems in the 21st century, Schneiderman (2000) argues that usability can be implemented in most of the household services and systems one utilizes. This concept of universal usability has increased the meaning of usability and tends to improve and provide guidelines for usability (Schneiderman, 2000). In today's highly digital society with emerging technology, usability can be implemented in various ways, and it is somewhat difficult to argue which technique or path is the right one to choose. One can argue that the choice of technique and evaluation

method is based on the situation, product or service that one wishes to conduct a usability test on.

2.4.1 International standards (ISO 9241-11)

As mentioned earlier, the International Organization for Standardization (ISO) defines usability as the "*extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use"* (ISO 9241-11, 2018). The ISO 9241-11, the standardization for usability, was inducted in early 1998 and has ever since been a widely accepted standard and highly implemented by many. The initial intention behind standardizing usability was to highlight the importance of usability as an outcome of interaction between man and product, software or service, rather than visualizing usability as an attribute of a product (Bevan, Carter, & Harker, 2015). Bevan et. al. (2015) argues that the emerging need for a standardization of usability has increased due to the vast use of the World Wide Web ever since its launch. Since usability is highly related to interaction between man and computer, the ISO 92491-11 performs an important role to minimize or eventually exclude issues and unwanted outcomes of interaction (Bevan et al., 2015). One of the key figures of the standardization of usability is that it can be directly related to both user and business requirements, where aspects such as effectiveness, efficiency and satisfaction can be used as measurements to achieve goals and determine whether users endorse the willingness to use a specific system (Bevan et al., 2015). One has to realize that it is urgent to identify the goals that need to be achieved in order to measure usability. Bevan et al. (2015) argue that the usability quality components can be used in various ways to measure, for example, satisfaction and performance efficiency of a product, generally in every situation of the product's or service's life span. However, a quality plan, that can be followed and documented, is needed in order to identify the needed level of usability for a product or service (Bevan et al., 2015).

Ever since the launch of ISO 9241-11, the standardization has been referred to by a vast number of studies and received much attention, as well as accompanied a great understanding towards usability and its importance. According to Bevan et al. (2015), the ISO 9241-11 spread the awareness of usability and even successfully acknowledged understanding among firms and organizations to adopt it in an early stage. However, the standard behind usability has been revised after its initial launch back in 1998, and since 2010, it has shifted more towards *Human-centered design,* but still acknowledges the same quality components such as effectiveness, efficiency and satisfaction, (Bevan et al., 2015). Bevan et al. (2015) note that one of the main differences in the revision of ISO 9241-11 is that components related to user satisfaction should include human-centered design requirements such as user experience.

During the past few years, the guidance on usability has furthermore been revised and today it is known as ISO 9241-11:2018, which replaces the initial version of the ISO 9241-11:1998. The main noticeable difference in the newly revised edition is the scope of usability, wider range of goals and clarification of quality components. The scope of usability includes other ISO standards coped with systems and services, such as ISO 9241-210, ISO 26800 and ISO 20282. The ISO 9241-210 is an international standard for *Human-centered design* for interactive systems with principles and guidance through the life cycle of a computer-based interactive system. ISO 26800 consists of specific concepts towards the usage of tasks and tools in order to provide safety and performance throughout the life cycle of a product or service. The ISO 20282 sets requirements for the context of use for everyday products and services.

The revised wide range of goals highlights the importance of both personal and organizational outcomes. Quality components such as efficiency and satisfaction have, furthermore, been revised and clarified within the revised ISO 9241-11:2018. With the revised standardization, the quality component of efficiency has shifted from defining accuracy and completeness among user task completion, towards focusing on efficiency in relation to the achieved results among users. The quality

component of satisfaction includes the importance and characterization of issues that can harm the user satisfaction. (ISO 9241-11, 2018)

The definition of ISO 9241-11 is generally widely accepted and followed up by a vast number of studies, however, the definition of usability by the ISO standard does not differentiate tremendously from Nielsen's point of view. Research in the field of usability, such as Nielsen (2010), defines usefulness and divides it into utility and usability, where utility answers the question of whether a specific system meets the need of the user at all. Usability in turn, focuses on how well the user can interact with the functionalities brought into the system. This can further be evaluated and measured by the five quality components seen in figure 6.

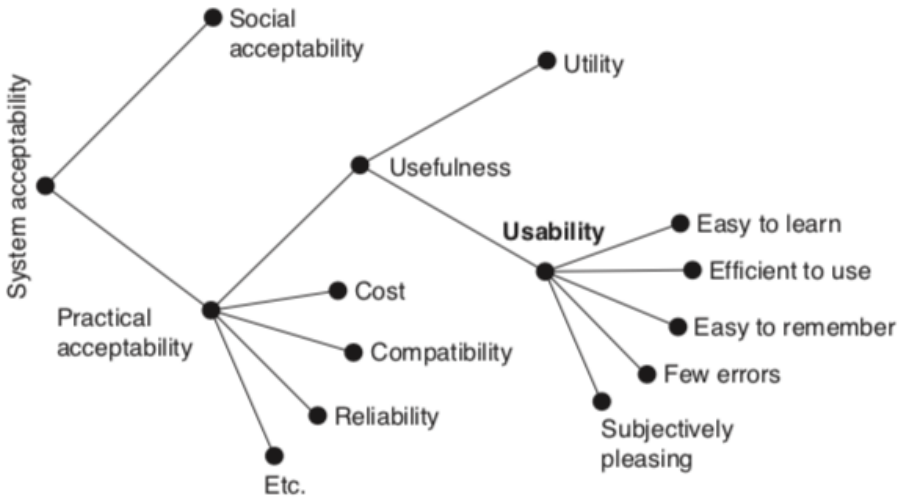


Figure 6: Different areas of usability in their context (Nielsen, 2010)

2.4.2 Importance of usability

The importance of usability cannot be highlighted enough. According to several researchers, usability among software production has been widely studied, resulting in the fact that it has a tremendous effect on the software production cost. Return on investment (ROI) is a widely used indicator, which is a performance measurement used to evaluate the benefit or return of a specific investment. According to Marcus (2002), the use of ROI is commonly used in evaluating and

measuring the value of IT functions, and widely used by organizations. Statistics show that the implementation of usability methods in software production, in order to increase the usability, increases the user's productivity and reduces occurred errors (Marcus, 2002). The reduced errors benefit the organization and the product development in many ways, where for example the need of user training can eventually be minimized. Marcus (2002) further acknowledges the increased savings with an efficient usability, especially when design changes occur, and the development must take a step back towards an earlier stage of the production. These above-mentioned factors, and many more, affect all internal return on investment. Usability methods can in fact also increase the external return on investment, by increasing sales as well as providing shareholders with guidance in understanding the value of the product or the whole company in question. Marcus (2002) further highlights that each product should require usability tasks that are specified for each product. However due to a vast amount of research done, usability tasks should be implemented in every possible step along the life cycle of the product or service. (Marcus, 2002)

Other researchers, such as Nielsen (2012), points out the importance of usability as "*a necessary condition for survival*", especially on the Web, where currently the average time spent on usability is only 10 percent of the design budget and twice as much when it comes to internal design projects. This will, however, not guarantee the result one is hoping for, since there is not a specific standard on how much to invest and when to invest. The key aspect here is to constantly keep usability as a corner stone of the project, as well as invest in usability already at an early stage or during the whole life cycle. (Nielsen, 2012)

2.4.3 Usability among people with disabilities

Former research show evidence that the Web is primarily designed for general public, as they perceive the usability of a certain website three times higher compared to people with visual impairment (Huang, 2003). However, designing a

website in order to be usable by people with disabilities, does not mean a reduced amount of content. A website that holds an accessible design, commonly increases the perceived usability for all its users, regardless of disability (Huang, 2003).

Prior study concerning usability among disabled people, indicate that a significant amount of money and focus has been invested in web developers, while the focus on disabled users is considerably lower. Previous research identifies a lack of recognition of the user's need for assistive technology. Therefore, the various forms and degrees of disability can cause hindrances that web developers have neglected. In order to observe the underlying usability problems among disabled users, a separation of the different forms of disabilities is needed. Additionally, a distinction between the variety of available assistive technology is required in order to ensure usability. Undoubtedly, as usability focuses on the user-friendliness of an interface, prior research indicates that users with disabilities should be involved in the early stages of the development process, to increase the usability of the website. (Bąkała & Bąkała, 2014)

2.5 Different perspectives of measuring usability

This section provides an overview of popular usability inspection methods. Firstly, the section will examine existing usability methods and discuss the theory behind them. Secondly, the theory behind the methods will eventually assist in determining which method to pursue in different scenarios. Thirdly, the theory will further identify which method or methods are the most suitable for this study.

According to Jacob Nielsen, software inspection as well as usability inspection has been used to improve the code of internet platforms and for evaluating user interfaces since the early 1990's. According to Nielsen, there are in general four elemental methods for evaluating user interfaces; automatical, empirical, formal and informal methods. Firstly, automatic evaluation is a selection of automatic evaluation techniques where computers automatically run the code through the

system, in to identify problems in the user interface. Secondly, empirical evaluation is considered to be more user-centered and requires user testing of an interface. The empirical evaluation method is consistently referred as the most common and as the easiest method to conduct. Thirdly, formal evaluating methods of user interfaces gather specific formulas and models in order to calculate usability measures. Fourthly, according to Nielsen, informal evaluation methods are "*based on rules of thumb and the general skill and experience of the evaluators"*. However, the success behind informal evaluation is highly dependent on the evaluators' former experience and knowledge. (Nielsen, 1994)

2.5.1 Usability inspection methods

There is a variety of usability inspection methods for evaluating user interfaces and systems. All the methods share the same goal of attempting to identify usability problems and hindrances within the interface, while the interface is being evaluated and inspected by an evaluator. According to Cheng and Mustafa (2014), usability inspection methods can be considered as a process of a products quality control, where instead of identifying flaws in a product, hindrances and barriers are identified in the user interface. A variety of techniques are allowing researchers and usability experts to conduct the methods in almost every stage of the project, however, according to Marcus (2002), in order to avoid expenses, the methods and techniques should be implemented at an early stage. Nevertheless, it is important to realize that usability inspection methods and user testing are not the same. Usability inspection methods can in fact be conducted without real user and do not require physical user in order to find usability problems (Cheng & Mustafa, 2014). The methods are, however, important to address in order to receive a wider perspective of available usability methods. Figure 7 illustrates the differences between usability inspection methods and the test approach, also known as user testing.

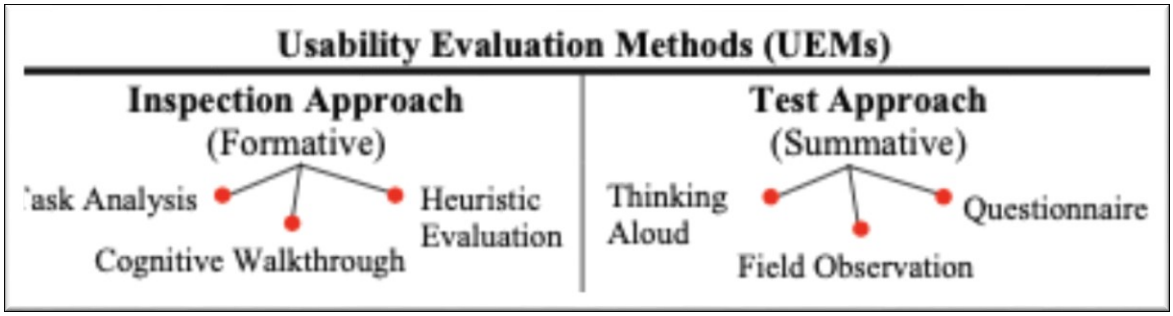


Figure 7: Categorizing usability evaluation methods (Cheng & Mustafa, 2014)

In the upcoming sections, the theory behind commonly used methods such as heuristic evaluation and cognitive walkthrough will be discussed. The decision to discuss these methods is because they can be classified as the commonly used usability inspection methods and should therefore be addressed. The corner stones of both inspection methods will be discussed as well as strengths and weaknesses. Additionally, the differences between automated usability testing and user testing will be shortly discussed.

2.5.2 Heuristic evaluation

The heuristic method for evaluating user interfaces was initially presented by Nielsen, in mid 1990's (Sutcliffe & Gault, 2004). The heuristic evaluation method consists of a set of heuristics, which are aimed to identify usability issues within an interface by using the heuristics as principles to find positive as well as negative aspects (Sutcliffe & Gault, 2004). The heuristic evaluation method is an informal method and one of the most widely used usability inspection methods in iterative design. The heuristic evaluation methods is conducted by examining an interface by one or several evaluators, preferably usability experts, while the evaluators are asked to comment and give their honest opinion about the interface (Atkins, Bennett, Domit, & Jones, 2011).

A set of heuristics, also known as guidelines or principles, are used by the evaluators in order to guide them through the interface and categorize the severity

of identified problems. The commonly used heuristics consist of ten general principles for interface design. Nielsen's 10 usability heuristics can be viewed in figure 8. The classification of the severity is performed on a scale of 1 to 4, where 1 indicates a problem that is only cosmetic (which is classified as a non-severe problem that can easily be eliminated) and where 4 is considered as a severe usability problem and crucial to eliminate. However, heuristic evaluation can be conducted by a single evaluator but in order to identify more usability problems, involving several evaluators is considerable. Pietre and Bevan further highlights the importance of defining the characteristics of the evaluators, since heuristic evaluation can be conducted by both novices an experts trained in usability methods. Nevertheless, experienced evaluators and usability experts are eventually able to identify more usability problems and determine the severity of the observed issues. (Petrie & Bevan, 2009)

Consequently, former study concerning heuristic evaluation, indicate that individual evaluators are only able to identify approximately 20 % or up to 50 % of the usability problems. However, the percentage may vary depending on the evaluator's expertise. Therefore, it is considered to involve several evaluators in a group, preferably a group consisting of five evaluators. (Nielsen & Molich, 1990).

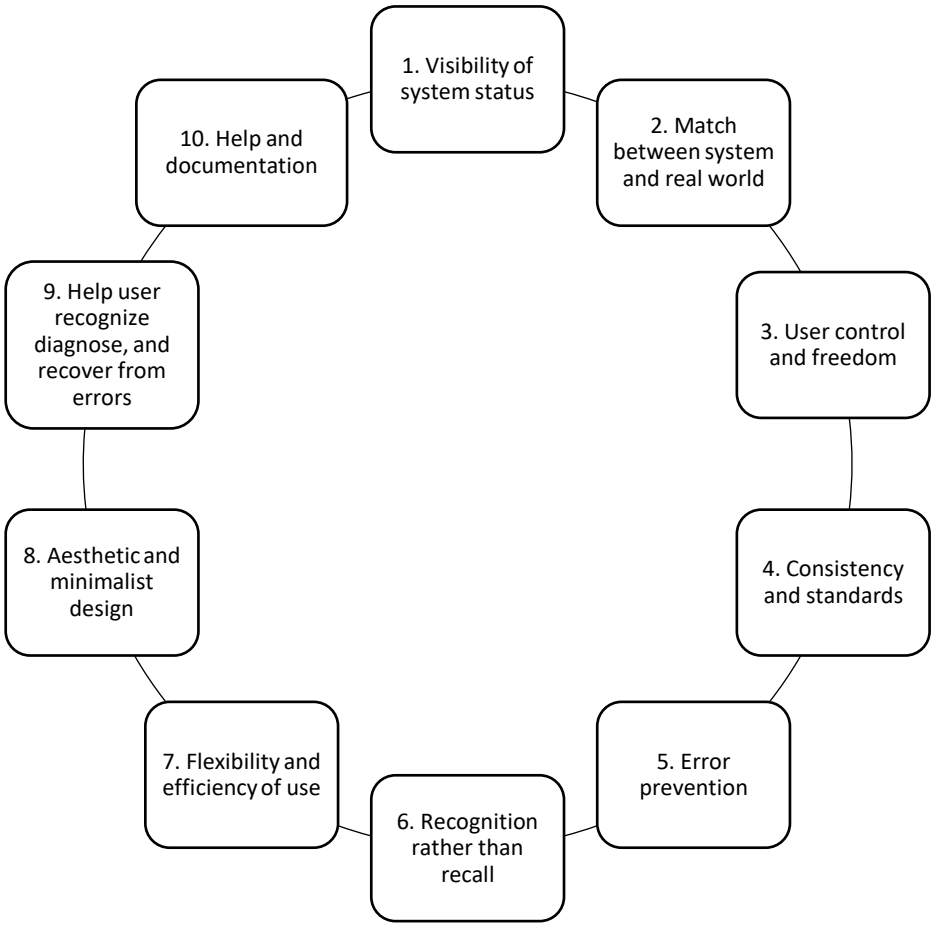


Figure 8: Usability heuristics for user interface design (Nielsen, 1994)

The heuristic evaluation method is considered to be a low-cost method that can eventually be applied as early as in the prototyping stage of an interface. According to Cheng and Mustafa (2014), the heuristic evaluation method can be combined with other inspection methods, such as cognitive walkthrough, in order to receive a wider perspective of the problems. However, the heuristic evaluation method as well as other inspection methods, do not require or involve real users and can therefore in some cases overlook certain problems (Cheng & Mustafa, 2014). Earlier studies identify heuristic evaluation as a usability inspection method that does not involve physical users, unlike usability testing where issues are found by observing real users communicating with the interface (Zhang, Basili, & Shneiderman, 1999). Heuristic evaluation rather identifies usability problems where evaluators review the user interface using various inspection tools and techniques.

2.5.3 Cognitive Walkthrough

Similarly, Cognitive Walkthrough is a usability inspection method that does not require real users to identify usability problems in an interface. According to prior research, the Cognitive Walkthrough is perceived as one of the most popular usability inspection methods, where the focus lies in evaluating how users act when interacting with parts of the interface rather than with entire interface. The purpose of Cognitive Walkthrough is to evaluate the learnability and the understandability of an interface by studying the mental processes of the participants. Cognitive Walkthrough is tasked based and the goal is to identify usability problems by creating real-life tasks reflecting the user's everyday interaction with the interface. Real-life tasks are created in order to address certain issues and to characterize whether the interface possesses enough guidance for the user to navigate throughout the system without major hindrances. (Cheng & Mustafa, 2014)

Although Cognitive Walkthrough does not require real users, it can instead consist of one or several evaluators (commonly usability experts) who explore the interface from the user's perspective. However, the process initially starts with identifying the users of the system, where aspects such as former experience and knowledge in similar systems is considered as key figures. Once the users are identified, it is important to characterize specific tasks that users would commonly face when interacting with the system. Although Cognitive Walkthrough can be implemented at any stage of the development process, a functional prototype or an already existing product is required. (Wilson, 2014)

According to Cheng and Mustafa, a commonly used process for Cognitive Walkthrough consists of specific stages. Figure 9 illustrates the method used when evaluating the interface of an online book retailer. Tasks are commonly arranged from *simple* to *more complicated* tasks, in order to observe how the mental state among the participants evolves while moving towards more complicated tasks. For analytical purposes, each task is followed by a specific question. It is important that the questions are as simple as possible, in order to clearly define what was done

and why the user decided to perform in a specific manner. The negative answers, issues and number of failed tasks will be collected and finally identified as usability problems. (Cheng & Mustafa, 2014)

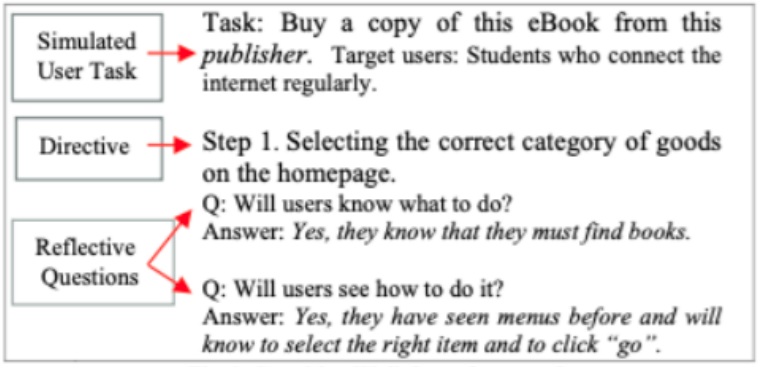


Figure 9: Cognitive Walkthrough method of an online book retailer (Cheng & Mustafa, 2014)

Prior research observes weaknesses regarding the narrow focus of the method, compared to other usability inspection methods. The narrow focus increases as the Cognitive Walkthrough is performed with an individual evaluator. As the focus of Cognitive Walkthrough lies in learnability and understandability, rather than in efficiency, the discovered usability problems are fairly low. Additionally, the process of the method and the short number of performed tasks, are considered to be other relevant weaknesses. Researchers suggest, that in order for Cognitive Walkthrough to be as cost-effective as possible, further methods and tasks as well as the pace of the whole process needs to be improved. Despite the negative aspects, there are also some strengths within the Cognitive Walkthrough. Firstly, Cognitive Walkthrough does not require a fully functional product or neither real users and can therefore be applied at all stages of the development process. Secondly, the method can be used by unexperienced users, however, usability experts or developers are preferred in order to discover a broader range of usability problems. Consequently, as Cognitive Walkthrough strictly focuses on cognitive theory and on usability attributes such as learnability, the method provides a deep understanding of how certain problems can be eventually eliminated. (Wilson, 2014)

2.5.4 User- vs automated usability testing

Usability is an umbrella term and consists of a variety of methods for testing a software's usability, however, several new methods in the field have risen and simultaneously gained popularity. Traditional usability testing and evaluation methods, involving evaluators and physical test participants, are still perceived as the most common methods, as the importance of the human factor is constantly increasing. According to prior research, the human factor is still vital and although the underlying code behind the website would be flawless, there is no certainty that the website is usable by a physical user (Rukshan & Baravalle, 2012). However, according to Rukshan and Baravalle (2012), the popular usability testing methods are time consuming and commonly require experts to perform and evaluating a test. Consequently, an increasing amount of money, time and other resources are invested in usability as technology is constantly evolving and new websites are rapidly emerging, the demand for usability specialists is expected to grow significantly (Rukshan & Baravalle, 2012). The increasing demand for usability analysts has resulted in several automated usability testing tools and also created guidelines for automated testing, excluding the human factor (Rukshan & Baravalle, 2012). According to prior research, it is expected that automated tools, used to identify usability and accessibility related problems, will continue to grow and increase in popularity as the need for mobile content enhances (Bader & Pagano, 2013). However, automated tools used to identify usability problems, are still at an early stage (especially among mobile applications). With the help of automated tools, researchers have been able to detect low discoverable usability problems, which Bader and Pagano (2013) describes as *"a specific usability issues type which occurs whenever a user interface does not communicate clearly that and how the user can interact with a particular element.* Noticeably, these are considered as problems that are overlooked by the commonly used usability methods (Bader & Pagano, 2013).

However, this study will overlook the use of automated tools, since usability problems ultimately depend on the end user and are therefore difficult to identify

using automated tools and methods. According to Norman and Panizzi (2006), a usability test should be user-centered and always involve both physical participants and observers. Although a usability test involving physical participants may in fact be more time consuming than an automated test, the possibility of receiving and documenting additional data increases as the human factor is present (Norman & Panizzi, 2006).

3. Company presentation: Kuntien Tiera Oy

Kuntien Tiera Oy is an in-house company owned by over 300 municipal organizations that produces ICT services exclusively for its owners. The company operates nationwide with offices in Helsinki, Kuusamo, Mikkeli, Tampere, Turku as well as in the Western Uusimaa. Kuntien Teria Oy provides productive high-quality ICT services to support the development of daily activities of its customers. The services offered by Kuntien Tiera Oy are designed to provide effective solutions and support for cities, municipalities and provinces all over Finland. Among the wide selection of ICT services provided by Kuntien Tiera Oy, this study will, however, focus on the service voucher and purchased service system, Palveluseteli- ja ostopalvelujärjestelmä (PSOP)

3.1 Introducing the PSOP system

Kuntien Tiera Oy, in cooperation with the municipals, has developed a system called *Parasta Palvelua* and its supporting PSOP system containing the effective management of service vouchers and purchasing services. The PSOP system is a nationwide information system developed by Espoo, Kouvola, Turku, Tampere and Oulu to support organizations and welfare services. The PSOP system is the most popular service voucher and purchase service system in Finland, where customers are able to view service vouchers and available service providers, both on the computer as well as on handheld devices. The PSOP system enables the customer to compare prices, quality and to choose the most suitable service provider for himself/herself. Furthermore, the system makes it easy for the customer to select a service based on the providers' location and on previous feedback from other customers. The PSOP system contributes with freedom of choice as a customer chooses between different services. Consequently, the use of the system is suitable independently of the municipality or service.

A service voucher is an alternative to self-produced or purchased social and health care services. An issued service vouchers can be used for social welfare and healthcare, such as personal help, domiciliary care, oral health care, and for many other services. However, the municipality determines the value of the service voucher and accepts the service providers whose services can be paid with the voucher. In addition, the customer can contact an approved service provider and pay all or a part of the service, depending on the value of the service voucher. Moreover, a purchasing service is a service that a municipality buys from a non-municipal entity and where purchasing services are used when the municipality does not carry out the service itself.

The following section will give an overview of how the PSOP system looks like and what it can be used for. It is important to observe that the interface of the PSOP system is the same on both mobile and on computer, however, the only difference is that the content on the mobile interface is cropped in order to fit the mobile device.

Three tasks will be performed in this study. However, in order obtain a clear image of the system, screenshots that introduce the various functionalities of the system are provided.
The first task consists of the service price comparison section, which is of significant importance for the whole system. The second task consists of service evaluation where customers are able to evaluate a recently received service. The third task consists of the feedback section of the PSOP system where customers can openly provide feedback to a service provider. It is important to point out that the information provided in the screenshots, concerning service providers and customers, are all fictional and do not contain any personal information. Screenshots regarding the task areas are illustrated in figures 17-25. Additionally, the task areas will further be discussed in chapter 5 and 6.

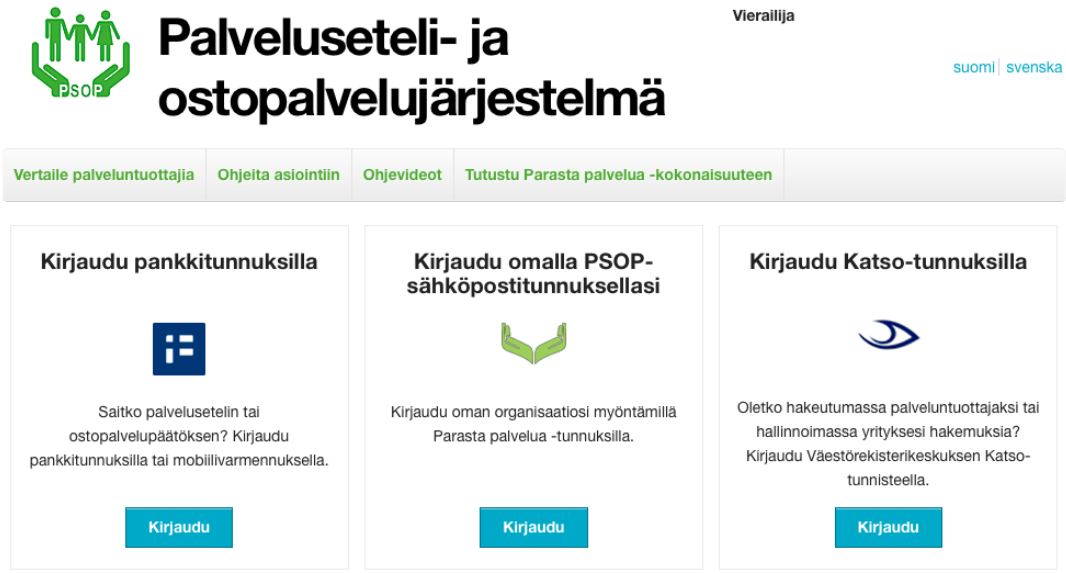


Figure 10: Login page of the PSOP system

Figure 10 illustrates the front page of the PSOP system where the user will arrive, when entering the site. Users are able to login to the system using different methods, either by authentication with bank codes or by using the PSOP email code provided by the organization. Additionally, the *Katso code* is intended for companies to identify themselves as service providers.

Figure 11: Customers can view personal information or choose to represent someone else

As the authentication is completed, users are able to choose if he or she wants to review the service vouchers and purchasing services for personal use or on behalf of someone else. The PSOP system also allows users to represent others who suffer from a disability or who are unable to manage the system themselves. However, in order to represent someone else, the user must be registered in advance as a representative. The person to be represented will appear in the drop-down menu shown in figure 11.

Figure 12: After choosing to view personal information, a personal front page will appear

Figure 12 illustrates a new home page that appears when the user has chosen to review personal information. Consequently, the user can choose to inspect basic customer information, purchasing power, purchasing balance, already provided services, provide feedback and eventually compare service providers in the service price comparison section.

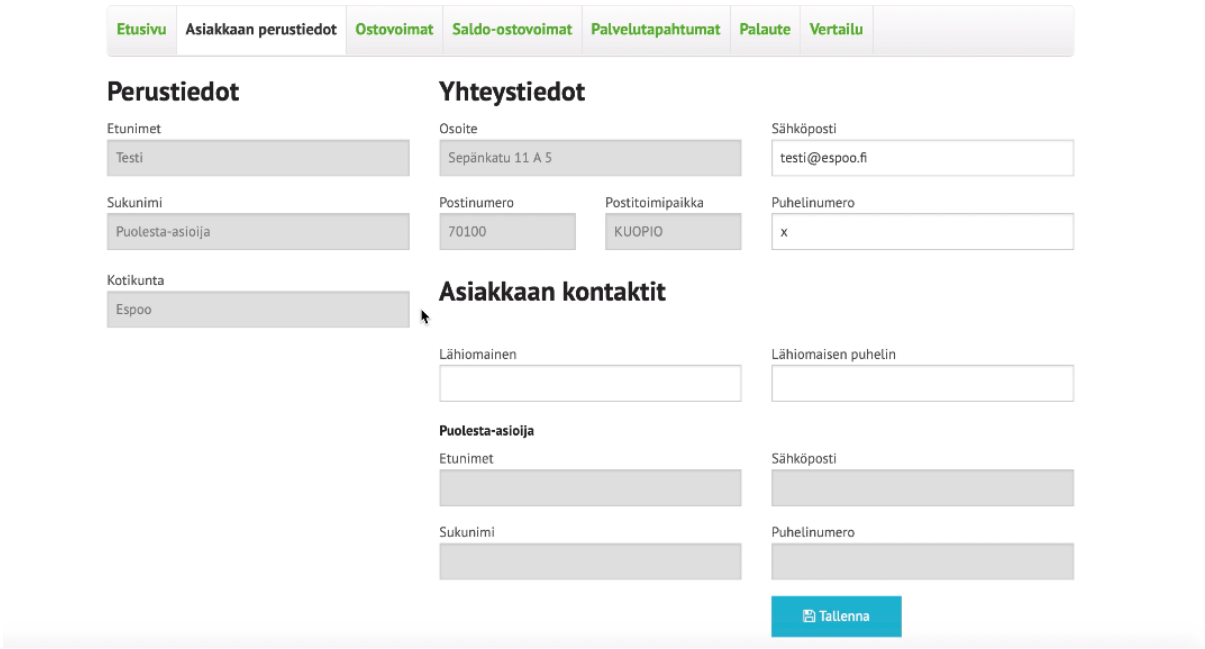


Figure 13: Management of the personal information

Additionally, customers are able change or add personal information, such as email address and phone number. The description of customer contacts such as a relative's name and phone number can also be added. However, the name and address of the customer cannot be changed, as they are predefined according to the authentication method used when entering the platform. Consequently, the columns that cannot be changed are illustrated in grey, as seen in figure 13.

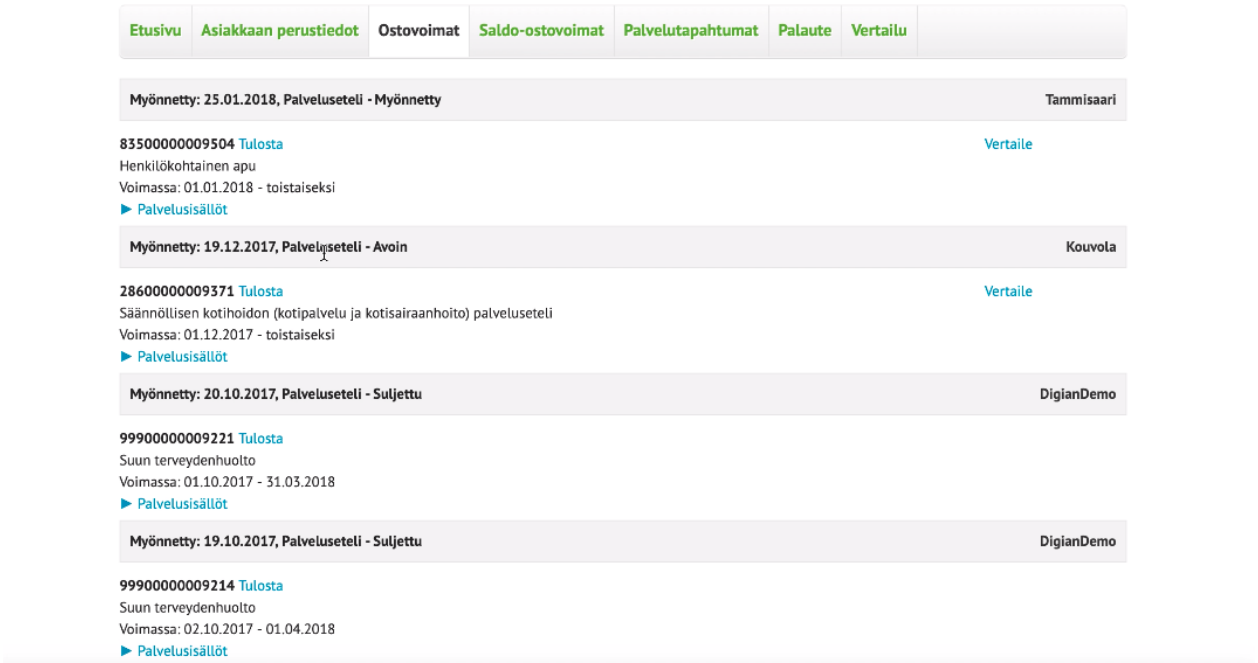


Figure 14: Customers are able to view their purchasing powers

When a customer receives a service voucher, the purchasing powers will appear on the following site, as seen in figure 14. The specified information regarding a purchasing power, can be accessed from the link labeled as "*palvelusisältö",* illustrated in blue. The link enables the customer to inspect the initial value and the remaining amount of the issued voucher.

A service voucher may be issued for a fixed or a predefined period, however, for a voucher that is valid until further notice, the end date is left blank. Additionally, the customer is able to view where and when the voucher has been issued as well as observe specified information of the service. Moreover, each issued service voucher has a unique number. Consequently, from this view, the customer is also able to navigate to the service price comparison to find both a suitable service and a service provider for himself/herself.

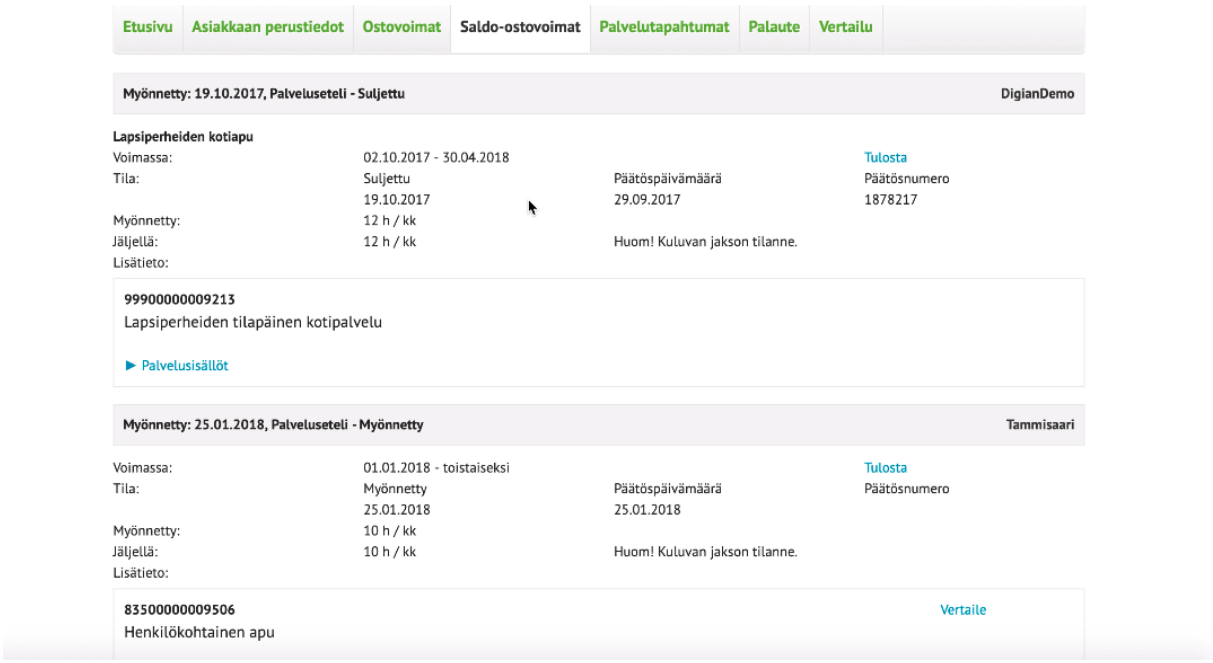


Figure 15: Customer are able to view the balance of the purchasing powers

The balance of the purchasing powers, including issued vouchers are illustrated in figure 15. However, the nature of the service voucher has to be selected as either unitary or euro denominated. Nevertheless, for unitary service vouchers, the unit, frequency, end date and a decision number will be issued. However, for euro denominated vouchers, the amount, frequency, end date and the decision number, are issued in the same manner. The interface operates in the same manner as the previously discussed purchasing power interface (figure 14), with the difference that these service vouchers have a predefined unit balance.

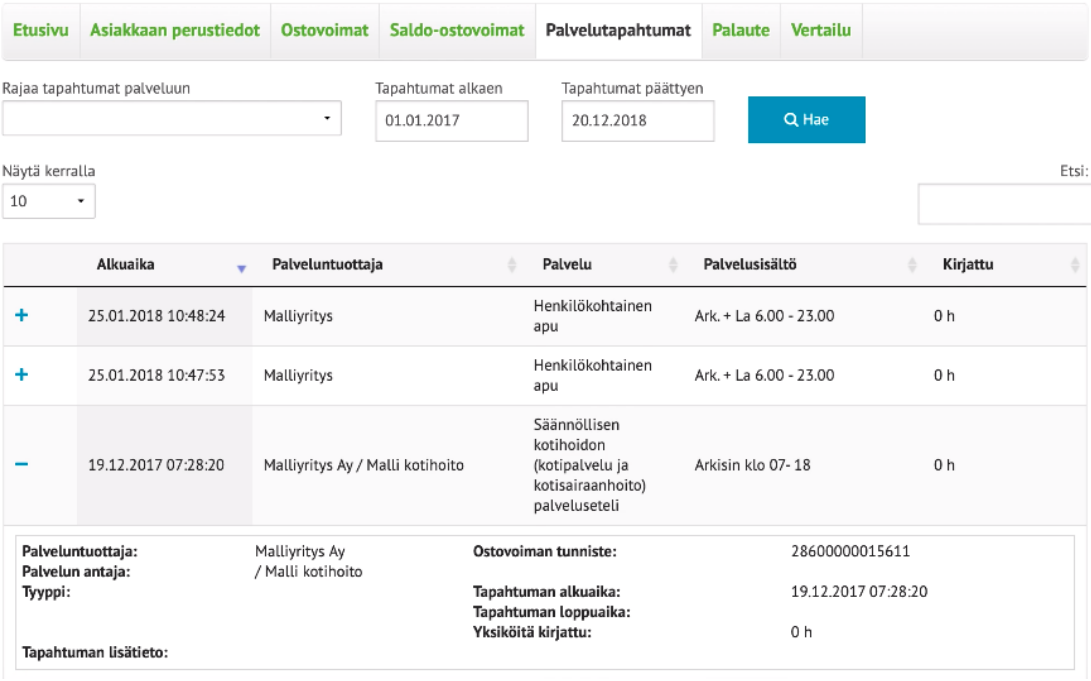


Figure 16: Customers can view further information about granted service events

Figure 16 illustrates which services have been granted to the customer during a specific period of time. On this page, the customer can limit the search results to a single service or according to a specific time slot, shown in the top left corner of figure 16. Additionally, a customer is able to freely search for a service event, which enables the customer to use desired search terms in order to find a desired service event.

In conclusion, the customer can view specified information about an issued service, such as the name of the service, starting date, content of the service and the number of hours registered. Interacting with the “plus sign”, on the left side of a row, provides detailed information.

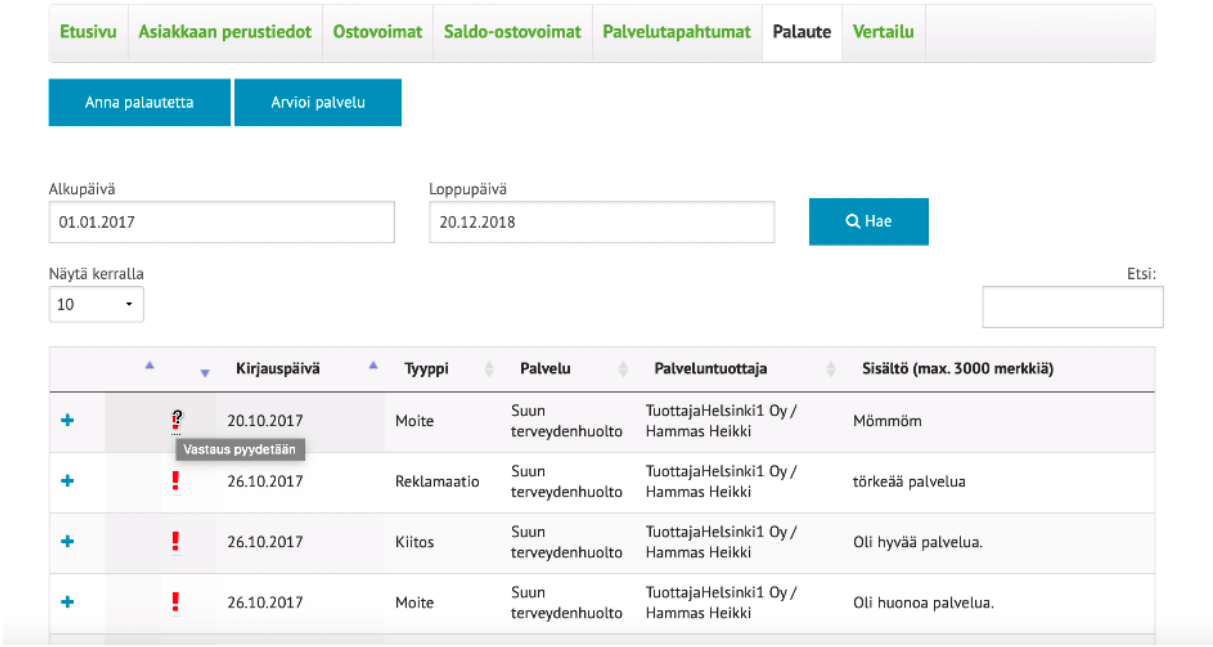


Figure 17: Feedback section, where the customer is able to provide feedback and evaluate services

Figure 17 illustrates the interface of the feedback section, which is a central part of this study. The interface shows the entry date of the feedback, type of feedback, which service the feedback is related to, service provider and the content of the feedback. In order to provide feedback, the customer accesses the link labeled as "*Anna palautetta*", seen in figure 17. Moreover, customers are able to provide the feedback to a service provider, direct the feedback towards a specified service and write a desirable message, as seen in figure 18.

Firstly, the service provider and the service to which the feedback is directed to is selected. Secondly, the customer can decide the subject of the feedback, as it can be labeled as "*thank you",* blame or as a complaint. Thirdly, the customer is able to write an open message in a desirable manner in the content section. Additionally, if the customer demands a response from the service producer, the customer is able to tick the box representing that a response is requested. Nonetheless, if the customer requests a response to the feedback, a red exclamation mark will appear in the table, labeled as "*Vastus pyydetään"* and illustrated in figure 17.

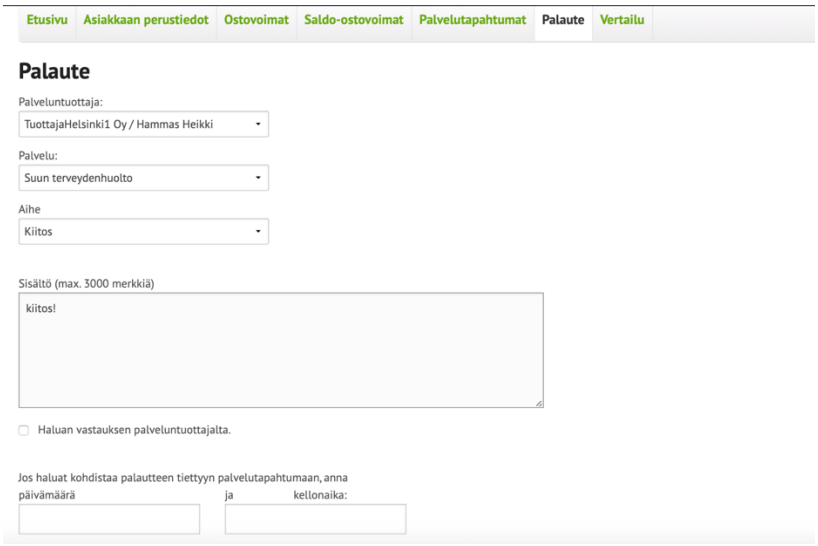


Figure 18: Providing feedback

The service evaluation form can be seen in figure 19. Customers are able to answer a short opinion poll, a so-called service evaluation, which is a query concerning the service provider and the service provided by the service provider. The service evaluation form can be accessed from the link labeled as "*Arvioi palvelu"*, seen in figure 17. The short query consists of 11 predefined questions and customers are asked to give an opinion between *completely disagree* or *completely agree*. Additionally, an option labeled as *"does not concern me"* is also available and should be selected if a specific question does not concern the customer. Consequently, a quality index, based on the evaluation form is calculated for the service provider. The quality index is displayed in the service price comparison section and illustrated as an image where the color of the image changes depending on the index, seen in figure 23.

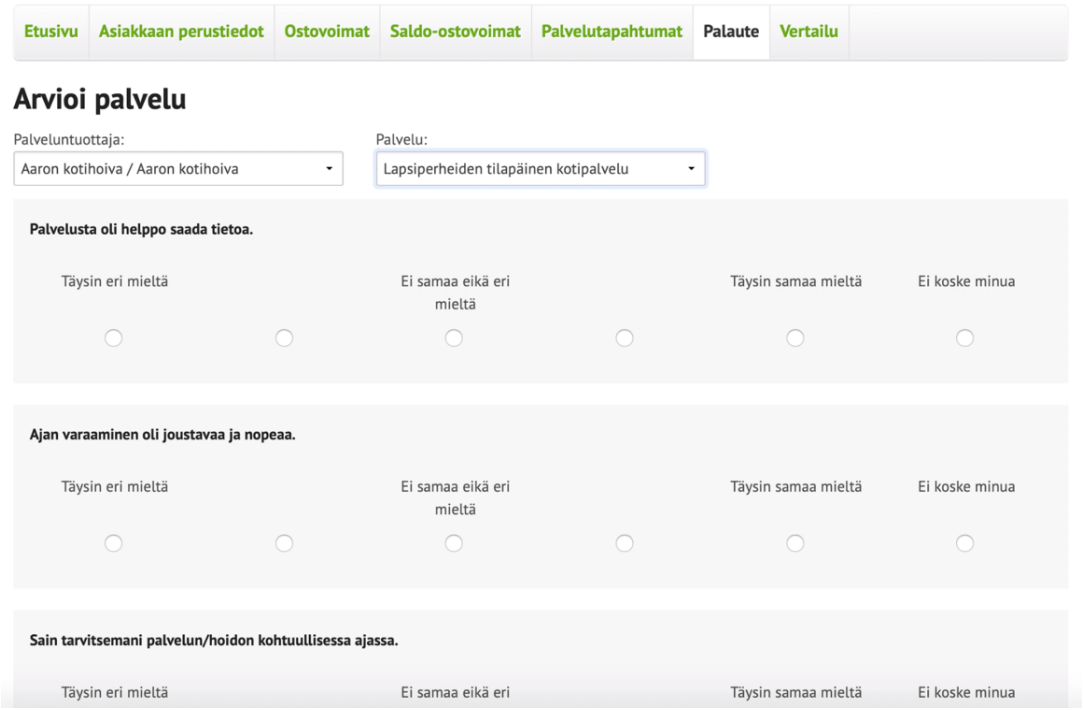


Figure 19: Service evaluation form

The home page of the service price comparison is illustrated in figure 20. Customers are able to find, distinguish and compare services provided by various service providers, both anonymously without authentication and as registered customers. In other words, this specific interface can be accessed if a user has not received a service voucher.

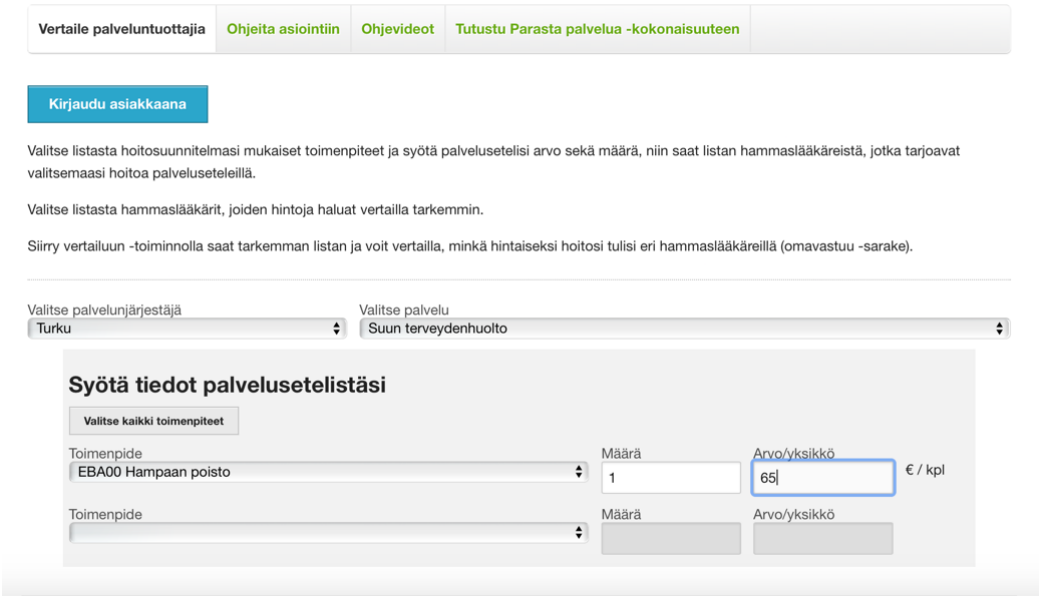


Figure 20: Service price comparison interface, where the customer is able to enter the details of the service voucher

In order to find a suitable service and a service provider, users must follow certain steps. Firstly, the customer should select a desired municipality after which the service category is selected. Secondly, the information of the service voucher is entered, as seen in figure 20. The customer specifies the procedure of the service and inscribes the quantity as well as the value of the voucher in corresponding sections. Thirdly, the customer is able to refine the search and specify search criteria in order to find a suitable service and a service provider, which can be seen on the left side of figure 21. Additionally, the customer can select a preferred language, identify themselves from the target group, define the accessibility and several other criteria as needed, seen in figure 22. However, the customer can also decide not to limit the search results at all, by ignoring the specification of the search and leaving all the checkboxes blank.

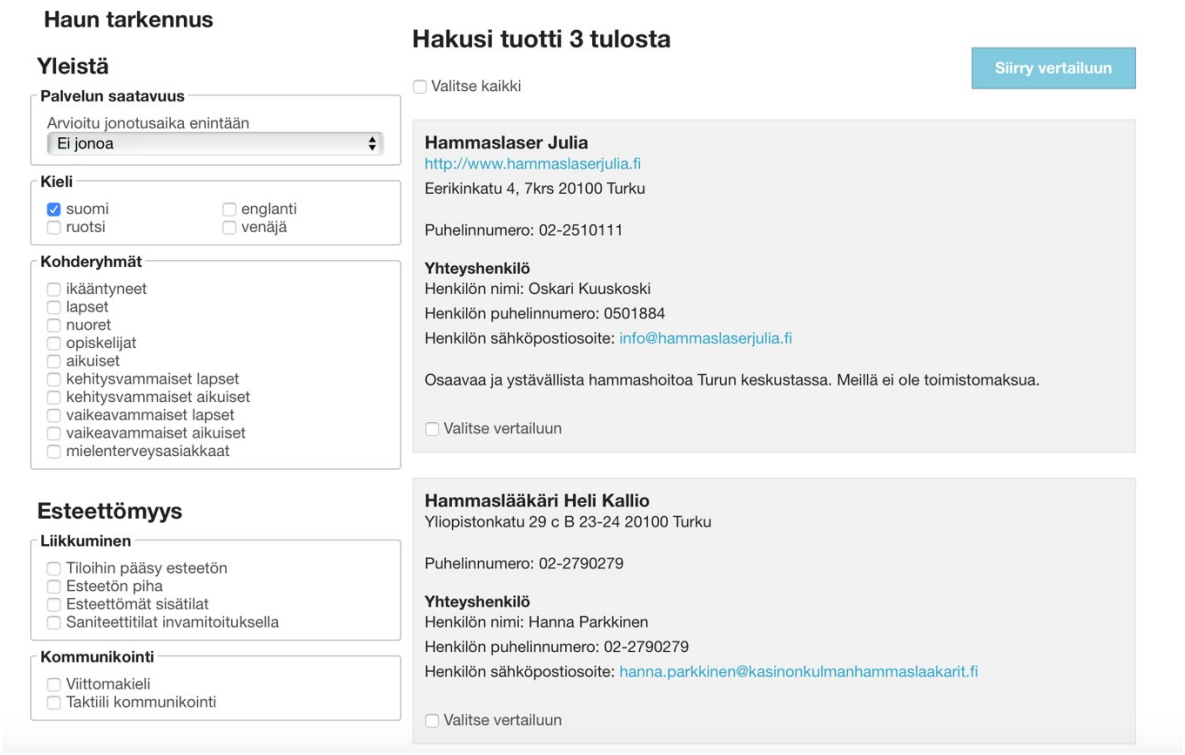


Figure 21: Detailed search and selecting service providers for comparison

The search automatically generates a certain number of service providers depending on the customer's criteria, illustrated on the right side of figure 21. The customer can select a desired number of service providers and proceed to the price comparison by accessing the link, labeled as *"Siirry vertailuun".* However, the customer is able to select all service providers to the comparison by ticking the box *"valitse kaikki",* before proceeding to the comparison. This will, however, result in a long list of service providers providing the desired service, which may or may not help the customer finding the optimal service provider.

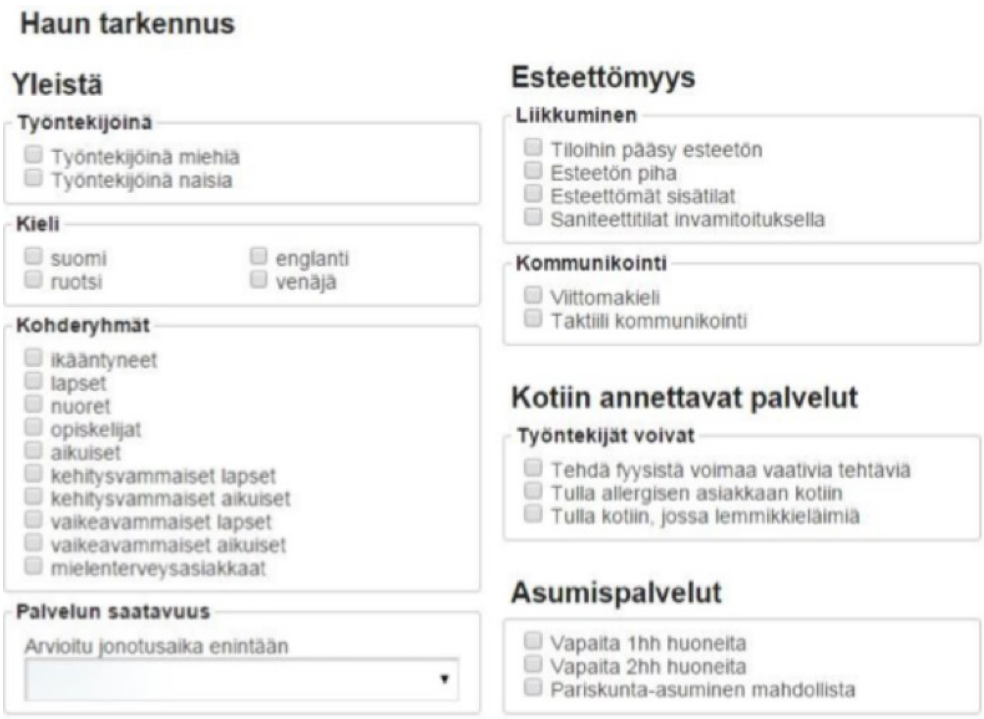


Figure 22: Further detailed search criteria

Since the customer has the opportunity to proceed with one or several service providers to the comparison, the interface may vary depending on the number of service providers selected. Figure 23 illustrates the final interface of the service price comparison. In other words, a new interface will open up and the customer is able to view the service providers as well as the deductible prices.

Figure 23: Service providers and deductible prices

However, if a service voucher has been granted through the PSOP system, the information of the service voucher is automatically entered, and the price displayed is the remaining amount to be covered by the customer. The deductible amount is only displayed for users who have been granted a voucher through the PSOP system. Consequently, if the service provider's service value is zero, the customers' service voucher will cover the cost of the service in question. In other words, the deductible amount will result in 0 €. However, if the service content of the service provider is empty, the service provider does not provide any service content for the service.

Additionally, the quality index is illustrated either as a question mark (if not know or calculated) or as an image with different shades of green color, see figure 23 and 24. Consequently, a service provider who has received enough positive feedback,

the quality index results in a smiling image in green, while service providers without feedback are left with a question mark.

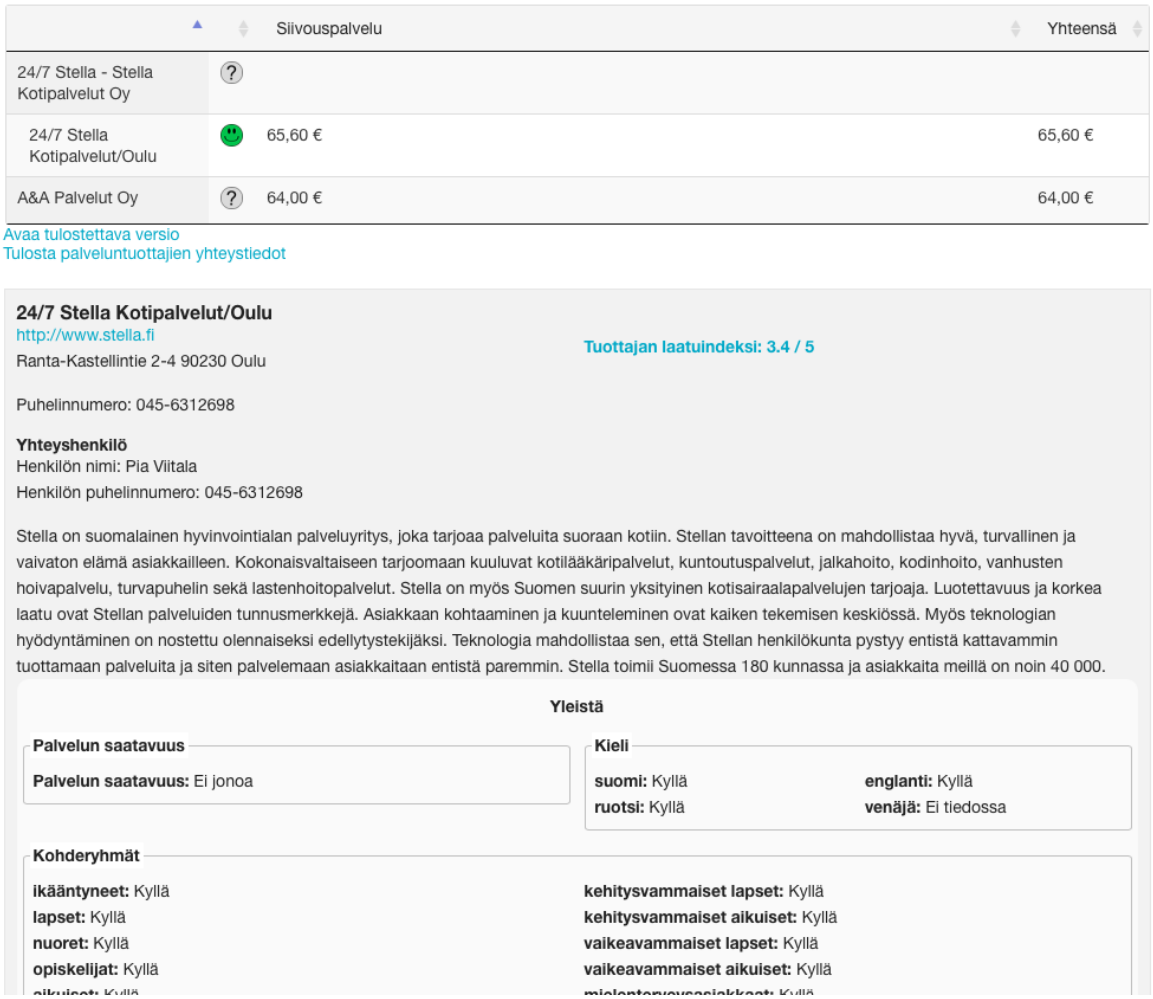


Figure 24: Additional information about the service provider regarding a desired service

According to the initial instruction, seen in figure 23, the customers is able to view additional information by pressing the name of the service provider. Additional information such as web page, address, phone number, quality index and a distinct description of the company is provided, see figure 24. Consequently, if the customer managed to find a desired service provider, he or she can now contact the company, as further information is provided.

4. Research methodology

There is an emerging difficulty in determining which method to be used when analyzing the usability and accessibility of a web service. According to research, it is important to identify whether it is more efficient to use methods related to usability or accessibility. When it comes to accessibility testing, several automated evaluation tools have been created in order to identify accessibility problems, however, these are more technical and, in many cases, excludes the human factor. According to the World Wide Web Consortium, involving the human factor in the methods is crucial due to the fact that a physical person is eventually the one to use the software. It is further mentioned, that no technical tool can and should not replace the interaction between the user and the software. Therefore, the World Wide Web consortium, provide methods for accessibility testing, where most of the content is identical to traditional usability testing methods. According the World Wide Web consortium, user testing is the most favorable method to visualize accessibility related problems due to the fact that it can detect problems immediately and can be conducted in various ways. However, according to prior research, it is suggested to involve a variety of real users with different disabilities in order to gain a wide insight of the systems accessibility. In addition, one of the key differences between usability and accessibility testing methods, is that the whole testing environment must to be fully accessible for the participants, regardless of disability. It is also noted, that the testing can and should be conducted in an environment that the user finds comfortable, i.e. at home of the participants. (W3C, 2014)

The methods that are used in this study will follow the methods of usability. The main reason behind this decision is that a more technical inspection of the PSOP systems accessibility issues have already been conducted. Additionally, it is not enough to design a web service to be accessible *"on paper"*, since it must be accessible and usable in reality by all its users. Prior study indicates, that this is a valid point to consider, as poor accessibility commonly correlates with poor

usability, and other way around (Mariger, 2019). Other researchers such as Brajnik (2008), indicate that although usability and accessibility are closely related, they have their own methods for inspection and evaluation. However, it is noted that when conducting a usability test involving people with disabilities, to identify and asses issues related to accessibility, all usability inspection methods can in fact be used (Brajnik, 2008).

The focus of this thesis will lay on empirical methods, as according to Nielsen, it is the main approach of evaluating interfaces and systems. Nielsen further points out, that difficulties can occur when recruiting participants, since in many situations it may result in unexpected expenses. According to other researchers and usability experts, only a handful of test participants are needed to identify the majority of the usability problems. However, prior studies identify that several inspection methods overlook usability problems but are also able to identify problems that other inspection methods may not find. (Nielsen, 1994)

Although section 2.4 discussed usability and gave an overview of usability in user interfaces, this section will further work as theoretical background and as corner stones that will support the methods used in this study to gather and examine data.

The objective is to identify the accessibility related issues that exist in the system by conducting a user test with physical participants. The purpose of the user test is to identify the main obstacles that people with disabilities face when interacting with the PSOP system. The method of user testing is the most appropriate, as it is fairly simple to conduct, as it involves real users and observers are able to gather data during the actual test session.

4.1 Qualitative and Quantitative methods for user testing

The method used in this study is a mix method between quantitative and qualitative user testing. The study will, however, lean more towards qualitative methods, since

participants with different disabilities are involved. As previously mentioned, it is three times as hard for people with disabilities to use the Web compared to for example sighted people (Huang, 2003). People with disabilities are also considered to be more time sensitive and therefore methods such as task-completion time, which is a common quantitative research method, may be left out of this study (Huang, 2003). According to Huang (2003), people with disabilities are also more likely to rate the severity of a problem significantly higher than non-disabled people. However, quantitative performance metrics such as success rate and number of errors, are aspects that will be examined (Budiu, 2017). However, according to Budiu (2017), problems may occur when conducting a user test with quantitative research methods, due to the fact that even if a participant would complete most of the tasks, it does not necessarily mean that the usability is good or bad. Furthermore, quantitative methods do not identify the characteristics of the occurred usability problems if the participant encountered hindrances and did not complete the task with a success rate of 100 (Budiu, 2017).

Qualitative methods, on the other hand, enable the observer to identify the characteristics of the problems that interfere with the task completion rate. Qualitative methods further allow the observer to characterize the difficulty of overcoming certain obstacles. However, the most important aspect of qualitative methods is that observers can interact with the participants and encourage them to speak their mind, resulting in further data analysis. A commonly used qualitative method is the *think-aloud* method, which will later be covered in section 2.4.2. (Budiu, 2017)

Consequently, a mixed method with of both quantitative and qualitative data, allows for collecting and analyzing data immediately during the user testing as well as after. According to Budiu (2017), qualitative methods are commonly used, but quantitative methods are the only metrics that eventually can clarify in numbers how the usability of a system has changed. Budiu (2017) further indicates that, qualitative methods are commonly conducted with a few participants in order to ensure flexibility. It is also noted that, approximately 85 percent of usability related

problems will be identified using qualitative methods and involving a small sample size of five participants (Budiu, 2017).

The purpose of involving both quantitative and qualitative research methods is to gain a better understanding and to further obtain a broader perspective of both usability and accessibility related problems. As a result, the mixed method also provides a better understanding of the users and their actions compared to using only one of the methods. In addition, the main reason behind using a mixed method for this study, is the fact that people with disabilities are involved. Because people with disabilities rate the severity of problems significantly higher and usually find it more difficult to interact on the Web compared to the general public, qualitative methods and metrics are considered to be suitable for this study.

4.2 User testing

User testing, also known as the test approach, is considered to be a combination of summative evaluation methods that require real user interaction in order to determine usability related problems with a system. According to Cheng and Mustafa, the corner stones of finding usability related issues lies in the test approach, as operates as the basis for the evaluator to receive both specific and contextualized feedback. Cheng and Mustafa further indicate that, it is important to recognize that usability methods changed and shaped over time. User testing, however, whether it is conducted in a lab or remotely, is still a method used more than half of the time in detecting usability problems. (Cheng & Mustafa, 2014)

User testing can be conducted in various ways, however, a number of participants are required to perform specified tasks or simply exploring their way through the interface (Nielsen, 2010). Although user testing might be considered as the most time-consuming evaluation method, it is surely one of the most cost-efficient methods to conduct (Nielsen, 2000). However, according to Nielsen (2000), people involved in designing a user interface, might have the perception that usability and

usability testing is expensive and eventually difficult to conduct, however, this is not the case. User testing can be conducted on the field, in a lab or even remotely, however, the environment should be pleasant in order for the user to feel comfortable. Undoubtedly, the importance of this aspect increases significantly when involving users with disabilities. Some researchers, however, highlight the fact that user testing should not be conducted remotely, due to the possibility of missing out on information (Brajnik, 2008).

In addition, the benefit of user testing is the ability to accurately map commonly confronted usability problems among real users, as usability problems may in fact have dreadful consequences. However, the severity increases when involving disabled users. (Brajnik, 2008)

4.2.1 Users

User testing starts with defining the users involved. The ideal scenario would be to involve users who are already familiar with the interface to be tested, resulting in reliable information. The users do not, however, need to be experts in using the product or service in order for the observer to gain valuable insight of the usability related problems. According to previous research, it is preferable to have a variety of users with a varying user experience, since novice and expert users usually face different types of usability problem. However, the number of users involved can vary and usually depends on the product or service to be tested. Prior research indicates, that a rule of thumb for user testing, is to involve no more than five users in order to find most of the usability problems.

Consequently, previous research by Nielsen, conforms this theory of only involving five users. Undoubtedly, it is obvious that conducting a user testing without users, results in zero observed usability problems. However, as soon as one participant is involved, and the data is examined, the observer can identify one-third of the interface's usability aspects (both positive and negative). When the second

participant is involved, there is a high possibility that the first and the second participant identify similar usability issues, as the use and interaction vary depending on previous knowledge and expertise. However, as users interact differently, the second participant may identify additional insights overlooked by the first participant. According to Nielsen, the third user, similarly finds the same insights as the first two, but on the other hand, the third user increases the findings with a fraction of new insights that were overlooked by the previous users. The identification of usability related problems will follow the same pattern until the fifth user is involved. In other words, the new insights will increase until the fifth participant, but only with a fraction. The additional new insights will, however, be reduced by involving a sixth participant, as involving new participants will result in similar findings. The pattern between involving participants in order to find usability problems is visualized in figure 25. (Nielsen, 2000)

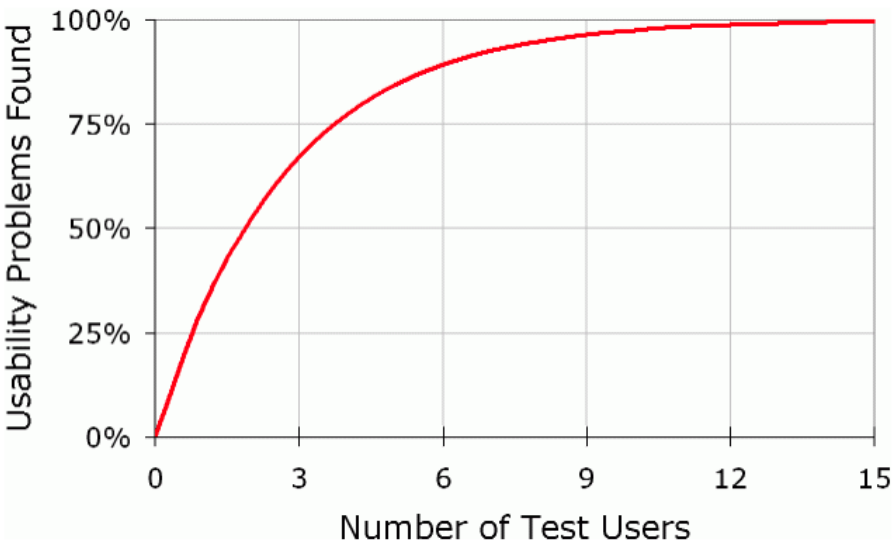


Figure 25: Proportion of usability problems found in correlation with the number of test users (Nielsen, 2000)

Additionally, Nielsen and Landauer confirm the theory of involving only five participants, due to the fact that the sixth participant would eventually find the same 85 percent of the usability related problems. Consequently, as only five users are required in order to identify approximately 85 percent of the usability related problems, different expenses can be avoided. (Cheng & Mustafa, 2014)

4.2.2 User testing with disabled users

As previously mentioned, approximately 15 percent of the world's population live with a disability of some kind (World Health Organization & World Bank, 2011). The number of people with disabilities is growing, as well as the ageing of the population which will continuously increase the number of chronic health conditions related to disabilities (World Health Organization & World Bank, 2011). Users with disabilities should and must today be included in user testing when aiming to identify usability and accessibility related problems, especially when the interface is provided by the public sector. (van der Geest, 2006).

However, when conducting a user test, involving people with different disabilities, several aspects must be considered. The whole test session, including tools, devices, assistive technology as well as the test environment itself, needs to be accessible from the very beginning (Hunter, 2010). Another valid aspect when considering the accessibility of the environment is to conduct the whole user testing at home of the participant. By doing this, the participant will feel more comfortable in his or her environment, while using the well-known personal technology. According to Hunter (2010), the fundamentals behind conducting a user testing involving people with disabilities, lies in having all necessary tools accessibly available. This includes screen readers, special keyboards as well as other types of assistive technology and tools.

According to former studies, people with disabilities have a higher possibility to experience hindrances when interacting with the Web. Disabled people are, furthermore, more likely to rate the severity of occurred problems significantly higher than non-disabled. Usability and accessibility related problems can be categorized in three ways; pure accessibility problems, which can be viewed as problems that only affect disabled people, pure usability problems, problems that only affects non-disabled people and, finally, universal problems which affect both groups. However, researchers are eager to point out the fact that problems occurring among disabled people and related to accessibility eventually affect the

non-disabled people as well. Designing the interface to address universal problems will be beneficial for both groups. Since usability- as well as accessibility problems affects both disabled and non-disabled people, researchers suggest that user testing, should in fact, be conducted with disabled users in order to efficiently identify and detect problems. Additionally, when conducting a user testing with disabled people, it is important to note that disabled people generally have a lower experience when interacting on the Web as well as with online services, compared to non-disabled people. The relation between universal problems and the aspect of user testing with disabled people, in order to receive a wider understanding of the overall problems has, however, not been widely studied. (Petrie & Kheir, Apr 29, 2007)

This is a valid point to analyze and an important aspect to cover, especially as the new law of Web accessibility affecting the web services of the public sector entered into force and will be gradually applied from the 23rd of September 2019 (Valtiovarainministeriö, 2018).

4.2.3 Recruiting the users

Prior researchers point out that only five participants are needed to discover the majority of the problems concerning an interface. Consequently, a careful recruitment process to find the right target group is needed. Since there will only be a small number of people conducting the user testing, the recruitment of the participants needs to be planned very carefully. According to Budiu (2017), there is a high possibility that the five participants may not, in fact, represent the target group, though the recruitment process has been carefully conducted. There is always a chance that the recruited participants do not represent the whole user population.

The fundamentals in recruiting the right participants is to understand who they are. One needs to clarify as much as possible about the required participants, i.e. who

they are, how they are related to this specific system and what benefits they have by using it. By doing this as early as possible, one is alleged to create more accurate profiles of the targeted user group. Nielsen Norman Group points out that understanding the study goals, and interviewing the stakeholders regarding their vision, will further ease the process and minimize the possibility of recruiting the wrong participants.

The recruitment process can either be internal or external, where external indicates that participants are needed to represent customers of a company or organization. For both processes, the recruitment processes can be conducted in various fields; at the company's facilities, universities, different agencies or similar. The key aspect is to carefully choose and contact the field according to the users, in order to save time and money as well as to maximize the possibility to recruit the participants as soon as possible. (Nielsen Norman Group, 2010)

4.2.4 Testing methods

An important aspect to consider when conducting a user test is to let the participant do the work and minimize the role of the observer. To minimize the role of the observer reduces the risk of affecting the mind and the opinion of the participants, as it may have an impact on the results. The involvement of the observer might also impact the task performance, due to the fact that the participant might try to please the observer's vision of the tasks or perform the tasks in a way that he or she thinks they should be performed. In order to identify the usability and accessibility problems among the participants and gain an honest opinion of the user interface, the role of the observer should be minimized. Methods such as the *think-aloud* process, is perceived as an effective method. The think-aloud process is a highly used method to gain further information and insights of the tasks as well as of the occurred obstacles. The observer has to clearly point out that this specific method is used, commonly at the beginning of the test session, in order for the participant to be aware of it. The think-aloud process allows the participant and the observer to have a short dialog whenever something

unexpected occurs. The process is of significant importance due to the fact that it allows the observer to identify and realize what the participant is thinking about, which could be left untold if the method would not be in use. It is important to realize that the think-aloud process does not, however, allow the observer to correct or guide the participant in any particular way. The process should only give the observer the image of what is going on in the mind of the participant. (Barnum, 2011)

According to Brajnik (2008), the think-aloud process among disabled people is especially vital, as the participants browse the interface with their specialized assistive technology, the observer should encourage the participants to think aloud. It is however noted that participants should interact with the assistive technology that they are familiar with, in order for the observer to address the appropriate issues (Brajnik, 2008). The think-aloud process enables the observer to simultaneously identify usability and accessibility related problems with the user interface. Brajnik (2008), however, points out that the role of the observer is to be minimized, similarly to an ordinary user test session, as it aims to ensure effectiveness and prevent the participants' actions and thought form being affected.

4.2.5 Turning interaction into tasks

In order to gain as much information as possible and to identify the severity of observed usability and accessibility problems, the participants should perform tasks on the interface instead of browsing their way through it. Only browsing their way through the interface, the observer may experience difficulties in identifying what is actually going on and might result in unwanted or undesired results. By predefining user tasks, the participants are able to proceed through the interface more systematically. A more systematic testing, including tasks, allows for a user testing that conducts the right parts of the interface and, furthermore, enables the observer to observe the right problems corresponding to the right segments of the

interface. Consequently, a task-based user testing is more realistic and further minimizes the possibility of participants being confused and elaborates the completion of defined tasks. (Moran, 2018)

Tasks should always be goal oriented and since the study will focus more on qualitative methods, the tasks need to be flexible and open-ended to ensure preferable outcomes. This is highly important among user testing involving people with disabilities. In order to conduct a task-based user testing, the goal that user should achieve needs to be identified before planning the session. The goal differs from interface to interface; however, the main idea is to perform activities and eventually use the interface for what it is meant to be used for. After the goal is clearly identified, one needs to figure out the different scenarios that the participant needs to go though in order to accomplish the goal. The importance is to not tell participant A to do *exactly this*, rather build up a short scenario and let the participant make his or her way towards the goal. This will enable the observer to evaluate *how* and *why* the participant decided to interact in a certain way. (Moran, 2018)

In addition, a clear description of the scenario should be presented to the participants, however, according to researchers, real life scenarios are recommended and highly desired. Furthermore, the scenarios should be as clear as possible to minimize confusion and further interaction for information between the participant and the observer. In conclusion, the purpose of building tasks, is to give a realistic view of how the interface should be interacted with. (Nielsen Norman Group, 2014)

However, it is important to realize that the think-aloud process is a highly effective method to be used in task-based scenarios. The observer should encourage the participant to speak his or her mind whenever facing an obstacle. This should, however, be done without giving away the correct answer to overcome the problem. Furthermore, in order to ensure the effectiveness of the whole test session, a pilot test is recommended to be conducted before the actual test. A pilot

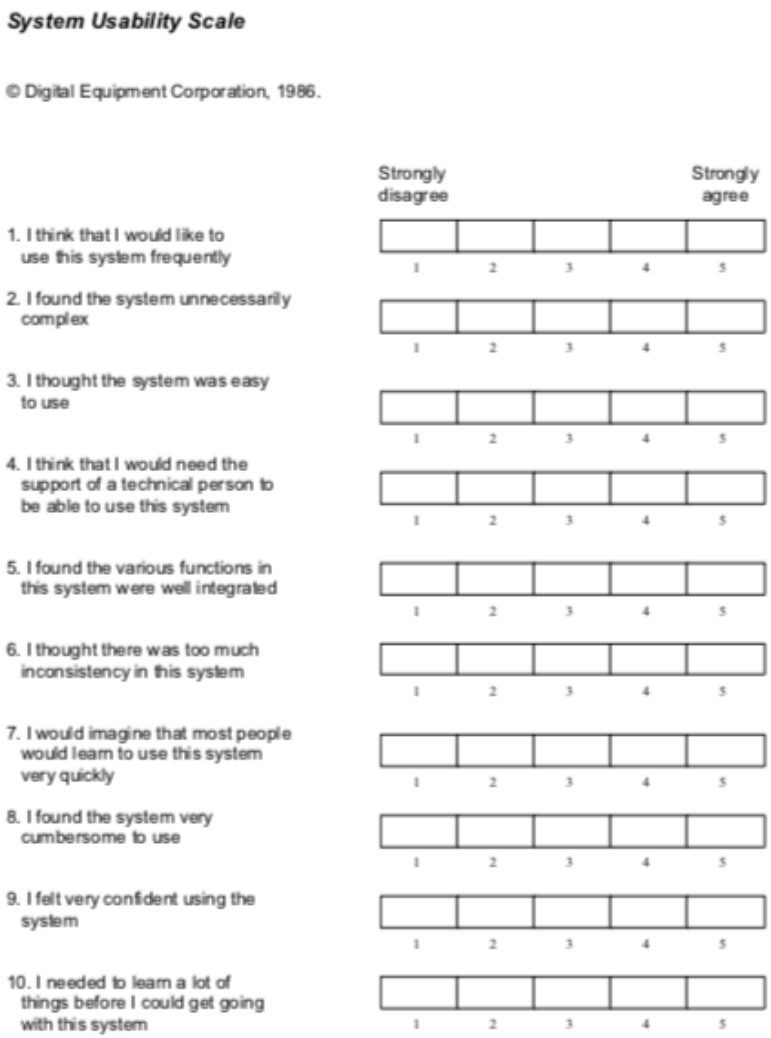
test enables the observer to build up the right tasks as well as correct any issues that might occur, before the first real test session with the target audience is conducted. The pilot test can be viewed as a practice session before the real test session, which allows for redesigning and tailoring of the whole user testing session, if needed. (Moran, 2018)

4.3 System Usability Scale (SUS)

There are numerous usability evaluation questionnaires and scales that aim to measure how users perceive the usability of a system. However, the System Usability scale, consisting of ten predefined questions and is considered to be the most established scale and a widely accepted measurement method. The System Usability Scale, also known as SUS, was originally developed in 1986 by John Brooke, however, later introduced to the general public in 1996. Since its launch in 1996, the SUS has received a significant amount of attention and its use has rapidly increased. The SUS was initially developed as a quick questionnaire for usability measurement purposes. Consequently, the SUS became a *quick and dirty* usability evaluation questionnaire that resulted in a simple and a trustworthy tool for measuring usability. Additionally, the SUS is based on the International Organization for Standardization (ISO 9241-11), regarding components such as effectiveness, efficiency and satisfaction, and earlier discussed in chapter 2.4.1. According to prior research, the purpose of creating the SUS was to establish a standardized questionnaire that allowed researchers to determine the perception of a systems usability. In addition, the questionnaire needed to be simple enough in order to be used in a relatively short time span. Although the SUS is considered to be a *quick and dirty* method for evaluating usability, the questionnaire has gone through a careful planning process in order to satisfy specific criteria. Firstly, a mutual connection between all the required elements was needed. Secondly, in order for the questionnaire to be simple and efficiently used, there needed to exist limitations regarding the elements. Thirdly, the questionnaire needed to be solid and follow a specific structure to make sense and not to be experienced as bizarre

by the users. As a result, a questionnaire consisting of ten predefined arguments, alternating positive and negative arguments and ranging from strongly disagree to strongly agree, was created. (Brooke, 2013)

The original SUS questionnaire can be seen in figure 26. However, in order to receive a reliable score, participants should always give their honest response and perception of the systems usability. Additionally, the center mark of the questionnaire should always be filled, if the participant is unable to return an answer on a specific statement. (Brooke, 1996)



System Usability Scale

© Digital Equipment Corporation, 1986.

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

Figure 26: System Usability Scale (Brooke, 1996)

The calculation of the SUS score is designed to be as simple as possible. Firstly, for all the positively oriented statements, statements 1,3,5,7 and 9, the score is calculated by subtracting one from the scale position (X-1). Secondly, for all the negatively oriented statements, statements 2,4,6,8 and 10, the score is calculated by subtracting the scale position from five (5-X). Thirdly, by summarizing all the numbers, a raw SUS score is calculated. Consequently, in order to calculate the final SUS score, the raw SUS score is multiplied by 2.5. Although the final SUS score ranges between 0 and 100, Brooke points out that the score must not be understood as percentage, which tends to occur. Additionally, according to prior research, it is necessary to grade the final SUS score in a visually efficient way, in order for people to understand it, regardless of their work experience or prior knowledge in the method. (Brooke, 2013)

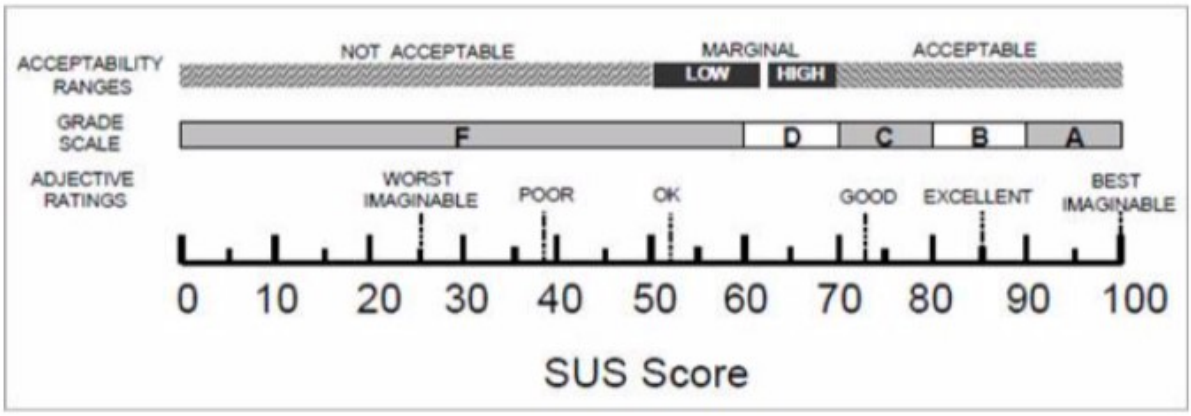


Figure 27: Grading of the SUS score (Brooke, 2013)

According to Brooke (2013), an efficient way to grade the SUS score is to visualize it from *worst imaginable* to *best imaginable*, as seen in figure 27. Figure 27 illustrates an efficient way to determine the acceptability, grade and adjective rating of the SUS score. For instance, all partners involved in a project can with ease observe that a score of e.g. 75 indicates that the usability of a system has a high and acceptable acceptability, a grade of C and an adjective rating slightly above good.

As previously stated, the SUS has been widely used and accepted, due to several of its advantages. Firstly, the small number of questions, with five answering

alternatives, results in an efficient way to measure usability and can be used without consuming unnecessary time. Secondly, the calculated SUS score is between 0 and 100, which makes it easy and convenient to analyze, even for managers, engineers and other non-experts in the field. Thirdly, the SUS is not tied to a specific field and can therefore be utilized to evaluate the usability of different systems, services and devices. Fourthly, it has been demonstrated that the SUS does not require a large sample size in order to retrieve reliable data. According to former research, a small sample of participants is enough to calculate a reliable SUS score. In addition, the SUS is not strictly tied to the English language, although it originates from it. Prior research show evidence that the SUS can in fact be translated into other languages and still prove to be as reliable as the English version. (Orfanou, Tselios, & Katsanos, 2015)

The SUS questionnaire used in this study was translated into Finnish and can be found in **appendix C.**

5. PSOP user testing

The previous chapter discussed the theory behind the research methodology that is used for this thesis. The theoretical background from the research methodology chapter works as the fundamentals for the whole test process. The following chapter will present the methods that are used in this study. Chapter 5.1 identifies how the target users were contacted and finally recruited. Chapter 5.2 presents the users that participated in the user testing. The chapter will present the characteristics behind the users and identify the reason why they fit the user group. Chapter 5.3 presents the necessary arrangements that were made in order to conduct the user testing. Chapter 5.4 presents the tools that were used, both the data gathering tools and the participants equipment will be discussed. Chapter 5.5 presents the actual test session of the user testing; the chapter will demonstrate the procedure of the test session and further present the tasks that were performed. Chapter 5.6 addresses the issues that were confronted during the whole user testing process.

5.1 Recruitment of participants

The recruitment process can be considered as the most challenging part of the study, since the target user group is relatively narrow. Firstly, the target users should all be blind and have knowledge in using assistive technology. Secondly, the target users should be users of the PSOP system and have received a service voucher through the PSOP system. Consequently, these two limitations eventually narrow down the target group significantly. In other words, the target user group consists of people who are completely blind using assistive technology and also clients of the PSOP system.

The recruitment process was significantly important and required a lot of patience as well as well-organized work. The recruitment process began in February 2019 by

contacting the city of Turku in order to reach the target group. In addition, a research permit had to be filled out to ensure that no laws or regulations were violated. However, the general data protection regulation was in the way of receiving any contact information from desired organizations or municipalities. Consequently, after weeks of hard work, the process of recruiting participants through the city of Turku, resulted in zero participants.

Undoubtedly, alternative procedures had to be done. Since the general data protection regulation limited the possibility of recruiting test users through municipalities and other public organizations, different recruitment processes had to be conducted. However, by personally contacting associations for visually impaired people, turned out to be the most efficient method. As a result, an *advertisement* encouraging people to participate in an important test, with the aim of improving everyday activity among people with disabilities, was forwarded to a mailing list for visually impaired people. The *advertisement* can be seen in **appendix A**. As a result, four people showed their interest in participating in the test. Since the *advertisement* was sent through a mailing list, no limitations regarding the location of the participants could be done. The majority of the participants acquired through the mailing list, were located in the Helsinki metropolitan area.

Additionally, as the minimum requirement for the user test was five participants and since the mailing list only contributed with four volunteers, additional actions had to be considered. Several emails were sent to a various organizations and associations without receiving any additional volunteers. Consequently, the remaining solution was to contact associations in person. A presentation of the thesis and of the entire project was held for the Southwest Finland Visually Impaired Association. The presentation was held for the people in charge and it took place at the facilities of the association. In addition, by help of the people in charge of the association, an effortless approach towards the target group was established. Moreover, the previously mentioned *advertisement* was also forwarded to the members of the association in order to address the importance of the test.

As a consequence of personally contacting the Southwest Finland Visually Impaired Association, four additional participants were recruited. The method of personally contacting organizations and associations turned out to be the most efficient and the least time-consuming method.

In conclusion, due to the fact that severe problems occurred during the recruitment process, several unplanned actions had to be made. As a result, participants were recruited in various ways, such as through mailing lists, municipalities, organizations as well as through in-person visits and presentations. The recruitment process resulted in receiving participants from Turku and the nearby Helsinki metropolitan area. A total of eight participants resulted from the recruitment process, which was sufficient since prior research advocate that only five test participants are needed in order to gain reliable data.

5.2 Participants

The carefully planned recruitment process enabled for finding the correct participants. The final user group consisted of eight participants, both male and female. In addition, the gender distribution can be seen in table 1, where the blue color represents male and red female. However, it is important to observe that both the pre- and post-test questionnaire as well the SUS questionnaire were created with Google Forms using the Finnish language. Therefore, some of the tables in the study are automatically generated in Finnish, which further makes it easier for Kuntien Tiera Oy to analyze.

The initial approach was to involve between 10 and 15 participants, in order to gain a large sample for the purpose of answering the research questions and to identify the usability and accessibility problems within the PSOP system. The user test was eventually conducted with a total of eight participants, following Nielsen's (2000) guidelines of only involving a small sample size in order to discover 85 percent of the problems. However, exceeding the recommended sample size of five

participant, did not interfere with the results, rather increased the belief and concept of Nielsen (2000).

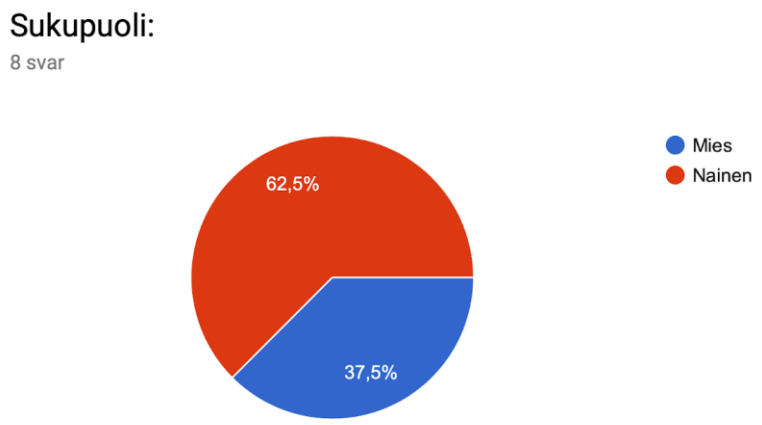


Table 1: The gender distribution between the participants

In order to fully answer the research questions, the participants needed to fit a certain user profile, which was well defined during the recruitment process. The recruitment process resulted in receiving participants who were all completely blind and used a variety of assistive technology, such as screen readers and braille keyboards. Since the target user group was fairly narrow, half of the participants were clients of the PSOP system, whereas the other half were not. The first half of the target group could test both the service price comparison, service evaluation and the feedback section. However, the other half of the participants were only able to test the service price comparison section of the system. This did not have a negative affect on the outcome, since the service price comparison section is of significant importance for the whole system and does not require a user profile or reception of a service voucher.

Consequently, all the participants rated their IT skills on a scale between one and five, where one resembles a *"beginner"* and five an *"experienced user"*. Surprisingly, all the participants rated their skills as three or above, identifying them as above average or as experienced users, seen in table 2. However, according to the participants, 62,5 percent (5 participants) preferred to browse the Web on a desktop computer or on a laptop, whereas the remaining 37,5 percent (3

participants) preferred the mobile platform. In addition, several participants noted that the assistive technology on mobile devices, especially the built-in screen reader on Apple devices, works effortlessly and is highly efficient. In other words, the mobile platform is commonly preferred and will further increase in popularity as ever more websites and applications become more accessible.

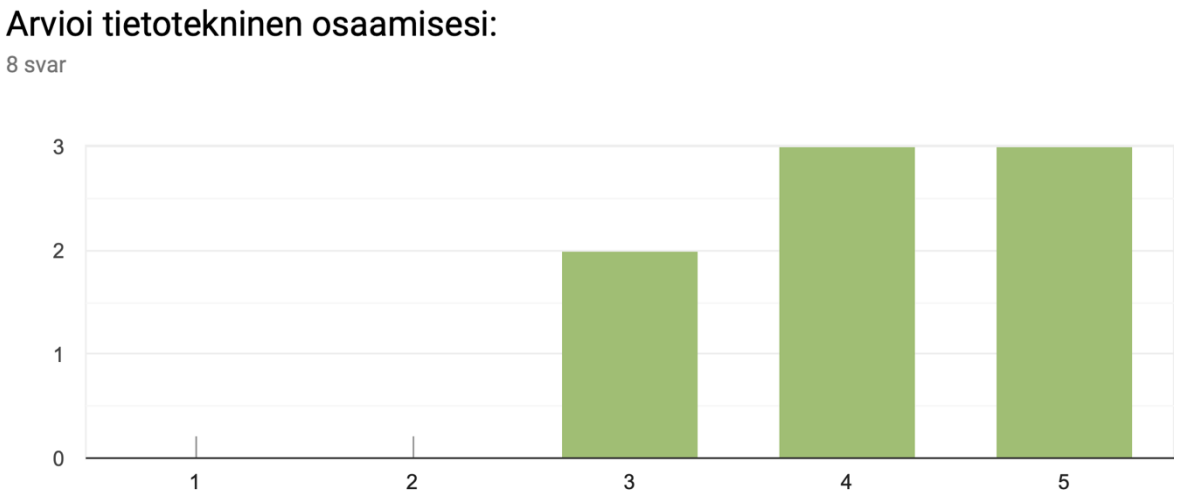


Table 2: Evaluation of the IT skills among the participants

In conclusion, users of all ages participated in the tests and the average age among the participants was 44,25 years. However, the age difference does not interfere with the results, since according to Eurostat (2015a), the usage of internet does not correlate with the age of the person, seen in figure 2.

5.3 Arrangements

The user testing required significant arrangements in order to ensure that the environment was comfortable and foremost accessible for the participants. Another valid aspect of the arrangements was to establish a neutral environment and resemble real-life interaction with the system. Consequently, as the user testing

was conducted both in Turku and in the Helsinki metropolitan area, the arrangements were done according to the participants' needs and requirements. The user testing, for participants located in the Helsinki metropolitan area, took place at the home of the participants and required further arrangements such as relocation and travel expenses. However, the other participants, who were located in Turku, preferred to conduct the user test at the facilities of the Southwest Finland Visually Impaired Association. Although the user tests were conducted at different locations, the arrangements were performed with an equal goal to ensure that the location of the test would not affect the results. Consequently, all the tests took place during the month of April.

5.4 Tools and equipment

The tools and equipment used during the user tests were rather simple, as even the most ordinary tools can be utilized to gather essential data. The author's tools consisted of a laptop and two cameras. The laptop was used to record the audio of the test session and to simultaneously follow the process of the test. In addition, with the permission from the participants, the two cameras were used to record the test session for further analytical purposes. The two cameras were used to record the facial expressions, screen and the use of assistive technology.

The tools and equipment of the participants varied depending on the user. However, all the participants used their own devices when conducting the tests. This was the most important aspect of the whole test session, since in order to gain reliable data, the participants had to use a familiar device with a recognizable assistive technology. Additionally, all the mobile devices were produced by Apple, using the built-in screen reading program *VoiceOver.* However, the setup for the computer varied slightly. All participants used screen reading technology supported by specified actions on the keyboard. Braille keyboards were also used, by a number of participants.

5.5 Test session

A user testing can be conducted in various ways using a variety of methods and techniques. However, it is important to consider that the tasks as well as the test itself, needs to undergo a carefully designed process in order to maximize the satisfaction of the goals. Another important aspect to recall, is the fact that the test environment has to be comfortable and accessible, especially when conducting user testing involving people with disabilities. These aspects were closely considered both during the planning process and when conducting the user testing. Consequently, the data was collected throughout the user tests, in order to uncover both usability and accessibility related problems. The data collection method consisted of task-based scenarios and of a pre- and post-test questionnaire.

The test session initially began with a pre-test questionnaire, which can be seen in **appendix B**. The pre-test questionnaire consisted of two parts. The first part gathered data regarding gender, age, disability and use of assistive technology. The second part gathered data of the participants IT knowledge, how often they use a computer or a similar device, number of hours spent on the Web per day, preferred device and how often they confront accessibility related problems while browsing the Web.

However, in order to ensure efficiency, the pre-test questionnaire was verbally presented to the participants and the corresponding answers were clearly documented by the moderator. This specific method was carefully planned and since all the participants were completely blind, a verbal presentation of the questions turned out to be the most efficient technique. Moreover, it was of significant importance to present the questionnaire verbally, as visually impaired people may experience difficulties filling out forms and lose interest in the test even before it starts.

The pre-test questionnaire followed up by presenting the tasks created by the author. A total of three tasks were conducted by the participants, see **appendix C**.

Task 1 consisted of the service price comparison, where participants had been granted a service voucher for tooth removal, by the city of Turku. Participants were asked to enter required criteria and select a desired number of service providers for comparison. The goal of the task was to find a company with the highest rating that provided the desired service. However, task 1 was completely fictional and did not include the use of a real service voucher. Task 2 was about service evaluation. Participants were asked to search for a service they had previously received and complete the service evaluation form in a desirable way. In addition, task 3 consisted of providing feedback to a received service. Participants were asked to browse through the system to find a desired service and finally provide feedback in a desirable way. However, the only criteria in task 3 was to ensure that the service provider will have to reply to the feedback.

A typical test session with a blind participant can be seen in figure 28. The anonymous participant in figure 28, uses both a screen reader and a braille keyboard. The figure further illustrates the tools used by the observer. This particular test session took place at the facilities of Southwest Finland Visually Impaired Association.

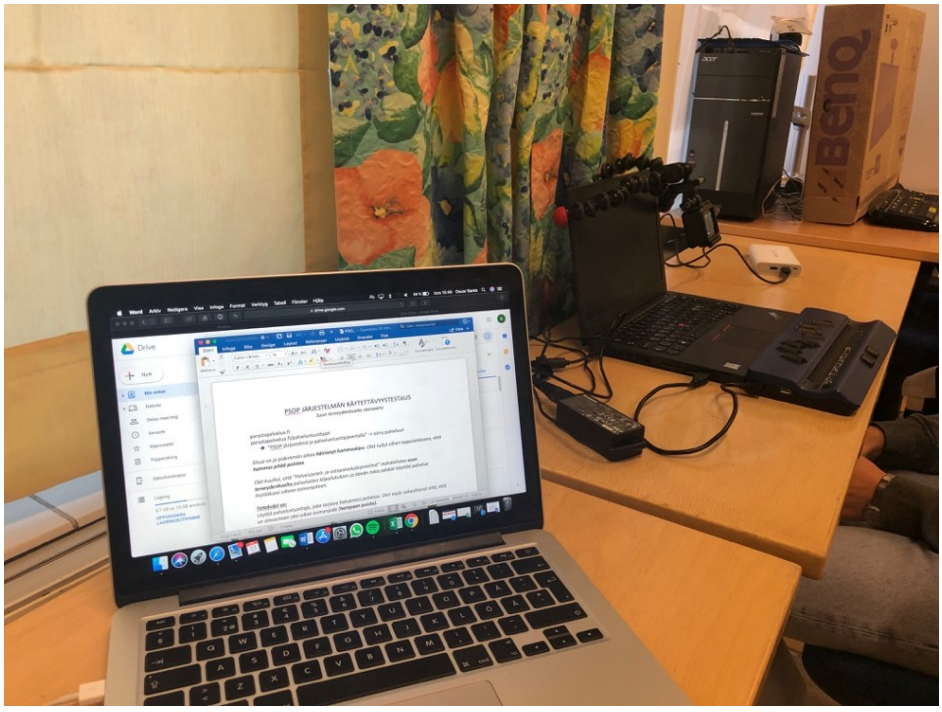


Figure 28: Test session with a blind participant using both braille keyboard and screen reader

It is important to note, that all participants were able to perform task 1, as it did not require a user profile nor reception of a service voucher through the PSOP system. On the contrary, only half of the participants were able to perform all three tasks.

Moreover, the user test of the PSOP system was conducted on both mobile and desktop platforms. However, the service price comparison (Task 1), was the only task conducted on both platforms with a total of eight participants. The other two tasks (Task 2 and Task 3) were only conducted on the desktop platform and involved 4 participants who had recently received a service voucher through the PSOP system. The variance of the findings between mobile and desktop platforms, is further presented in chapter 6.2.

Furthermore, the final phase of the test session consisted of a post-test questionnaire. The previously discussed System Usability Scale (SUS scale), consisting of 10 predefined questions, was used as the post-test questionnaire, see **appendix D**. The initial plan was to distribute the questionnaire through a mailing list to the customers of the PSOP system, however, the plan was brought to an end due to the General Data Protection Regulation. Nonetheless, the SUS questionnaire could effortlessly be used as a post-test questionnaire. Similarly, to the pre-test questionnaire, the questions and claims were verbally presented and noted by the moderator.

During the tests, conducted on both mobile and on computer, the participants performed tasks while the moderator simultaneously followed their progress. However, the duration of the tests varied slightly depending on the user and their IT knowledge. The think-aloud method was also used throughout the tests and participants were encouraged from the very beginning to speak their mind whenever something occurred. Additionally, questions were asked during and especially after every task, in order to gather as much data as possible. Further specifying questions were also asked, whenever the moderator expected that something had been left unsaid. Consequently, the purpose of this method was to

analyze even the slightest problems related to usability and accessibility, which might have been overlooked by only analyzing the recorded material.

As a result, the user test provided an enormous amount of data, as several hours of video and audio material was gathered. However, efficiency nor task completion time were analyzed, since according to prior research, task completion time is not a perceived metric to measure when people with disabilities are involved, as it might increase the possibility of terminating the whole test session. However, the number of usability and accessibility related problems and the severity of them were analyzed. Due to the fact that the role of the moderator and the evaluator was performed by a single person, the number of problems were not listed during the test, however, they were later analyzed from the recorded material. As a result, this method was chosen in order to maintain an open dialog with the participants and similarly minimize interruptions.

5.6 Issues

As a result, a tremendous amount of data had been collected from the user tests. Undoubtedly, the data had to be carefully organized and eventually duplicated for backup purposes. In addition, data gathered from the questionnaires were fairly simple to organize, as the early decision of using *Google Forms* turned out to be a highly efficient tool and did not require any major arrangements or manipulations.

The recorded material produced roughly six hours of both video and audio material, which equals to approximately 79 gigabytes. Both the video and audio material had to be transcribed, in order to ease the process of identifying problems. The average test time per participant was roughly 43 minutes and resulted in an extensive data analysis. Consequently, the transcription process identified both usability and accessibility related problems and enabled for analyzing the participants comments. In addition, possible solutions for the identified problems were documented, which

will be presented in chapter 8.2 where future recommendations for Kuntien Tiera Oy are discussed.

6. Results & analysis

The following chapter presents the results from the user tests. Chapter 6.1 provides an overview of the observed problems. The chapter will briefly discuss the overall problems and how they were analyzed. Additionally, task specific problems will be discussed. Chapter 6.2 will present the difference between the problems found on mobile and desktop platforms. Chapter 6.3 will present the findings from the survey, as calculations from the System Usability Scale is presented. Chapter 6.4 will discuss the methods and metrics that were used in order to evaluate the problems.

6.1 Findings from the user testing

The purpose of conducting a user test of the PSOP system, is to analyze and determine if usability and accessibility problems are present within the system. Another purpose is to provide valuable insight for Kuntien Tiera Oy, in order to take the problems into consideration when further development of the system is done after the second quarter of the year. Consequently, after examining the problems and errors within the PSOP system, it can be determined that the system does not fully comply with WCAG, as both usability and accessibility are still present. As a result, improvements are needed in order to completely satisfy the criteria that Valtiovarainministeriö (2018) requires for websites provided by the public sector.

As a result, both usability and accessibility related problems are present in all of the tasks that were conducted. However, the number of encountered problems and the severity of them varies depending on the user, their experience and the willingness to overcome these problems. Noticeably, a number of the problems were severe

enough for some users, making certain areas in the system hardly accessible. Undoubtedly, all participants experienced some difficulties and confronted various problems, however, by using the think-aloud method some obstacles were vanquished after participants truly concentrated and managed to find a way around a certain problem. Consequently, all participants experienced problems, and none managed to perform a task with a success rate of 100 percent.

Results were gathered and analyzed from all the methods used in the study. It can be clearly stated, that the user testing produced the majority of the findings, as a substantial amount of the encountered problems were found during the test session. Additionally, by using the think-aloud method during the test session, further meaningful insight from the participants could be observed and recorded. However, the majority of the problems were identified after analyzing the recorded material, as the recorded material enabled for further analysis and provided a valuable insight of the problems that were either forgotten or not noticed during the test session. Additionally, the pre-test questionnaire uncovered the characteristics of the participants and made it possible to analyze if the observed problems correlated with their IT skills. The post-test questionnaire consisting of the SUS, gathered opinions regarding the PSOP system, in order to calculate how usable and user friendly the system currently is. However, since the number of involved participants followed the guidelines of Nielsen (2000), the results from the SUS questionnaire were rather limited but possibly pointing to the right direction. The findings from the questionnaire will be further discussed in chapter 6.3.

In conclusion, the analysis of data shows that the PSOP system has both usability and accessibility related problems and can therefore, to some extent, be inaccessible for people with disabilities. Although a technical accessibility test of the PSOP system has been made, user testing shows evidence that problems and issues are still present. As a result, it can be clearly stated that although a system is technically accessible, a user testing can discover how accessible the system actually is and reveal the real issues among its users.

6.1.1 Findings from Task 1

Task 1 was about the service price comparison and the goal was to find a service provider that had received the highest rating. The task was divided into three parts. The first part of task 1 can be seen in figure 29. According to the task description, the city of Turku has granted one service voucher for tooth removal with a value of 65 euros. Participants were asked to find the correct municipality and service category from the dropdown menu. After choosing the municipality and the desired service, participants were asked to enter the information of the service voucher.

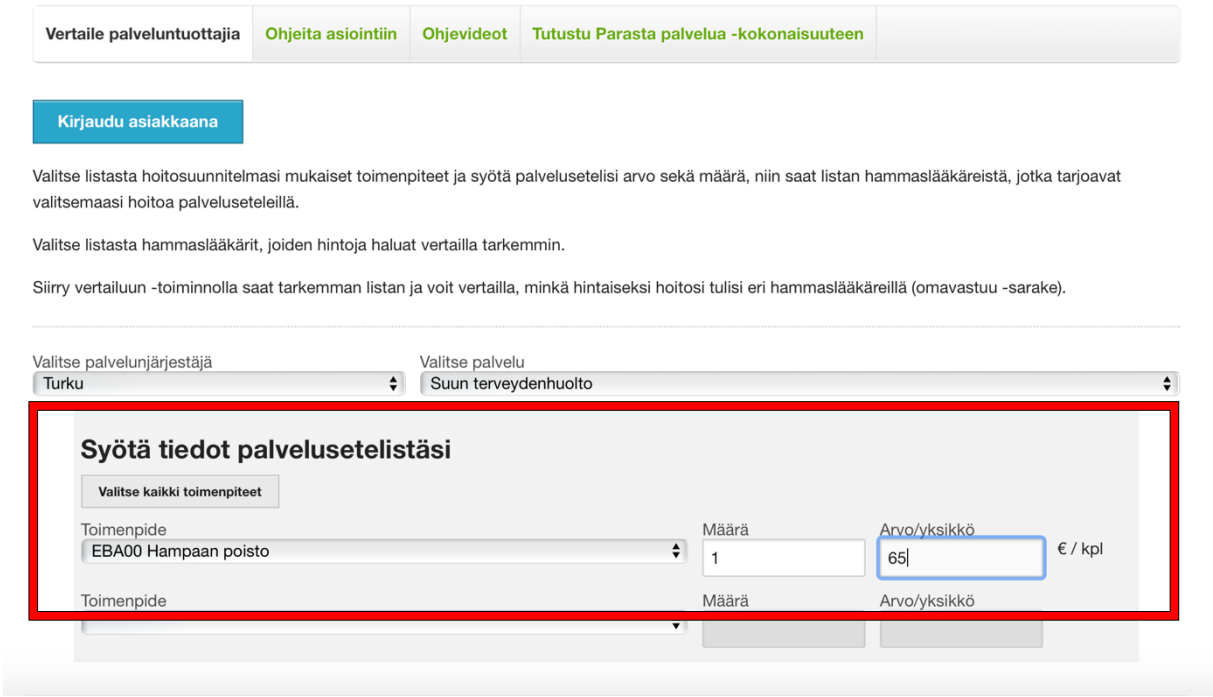


Figure 29: Part one of Task 1

Consequently, the participants did not experience any major problems with selecting the right municipality and service category. However, two participants noted that even though the right service category had been chosen (*Suun terveydenhuolto)*, they were not aware if it had been chosen correctly, since the assistive technology did not inform about it. According to one participant, this might cause problems and leave the user clueless, since the screen reader does not indicate if it has been correctly selected when returning to the section. Nonetheless, the same participant noted, that users might experience difficulties in

understanding the terminology used on the site, as the participant was not fully aware of the fact that *"palvelunjärjestäjä"* referred to the city or municipality that provides the service. Additionally, two participants pointed out that it would be more user friendly to implement the titles of the dropdown menu in the drop-down menu itself, instead of above them as they are now. However, this obstacle did not cause major problems and the participants managed to discover the meaning of the section on their own. Nevertheless, the participant noted that a simplified terminology would be necessary and less time consuming for blind people

On the contrary, severe problems started to occur when the information of the service voucher was entered. As a result, it can be clearly stated that entering the information of the service voucher was the most severe problem that the participants faced, as all eight participants experienced major problems with this section. The problem area is illustrated with a red square in figure 29. All eight participants had difficulties choosing the desired procedure of tooth removal and determine the quantity and the value of the service voucher. Consequently, two of the participants were unable to enter any information of the service voucher. Furthermore, participants noted that the system was inconsistent and severe problems occurred when the service voucher section did not follow the same structure as the earlier section. Two thirds of the participants pointed out that the titles were programmed in an inaccessible way, since the screen reader first read the titles and then read the empty boxes. All participants noted that this inaccessible programming of titles results in a so-called *guessing game*, which is a commonly encountered problem for people using assistive technology. The so-called *guessing game* is a problem that will appear frequently throughout the user testing, where participants have to remember and eventually guess the exact order of the titles in order to know which title refers to which box or section.

In addition, this specific section also left the participants clueless whether the proper procedure of tooth removal has been selected, since the system or the assistive technology does not indicate it in any way. Participants were also eager to note that they were totally unaware of the fact whether the right quantity or value

had been entered to the proper box. Consequently, as participants managed to fill in the information of the service voucher, a new section opened up below (enabling users to add further actions) and confused the participants. However, some participants understood the meaning of the new section and were able to identify what happened and erase the unnecessary information, while other participants maintained stuck and confused.
In conclusion, all participants noted that this specific section of the system was inaccessible and not user friendly at all, as it creates frustration among users. According to the participants, it is also highly time consuming to go back and forth, trying to identify if the right section has been entered correctly or not. Nonetheless, two participants pointed out that this is a common problem in most online services and can easily be reprogrammed by implementing the titles into the corresponding boxes.

According to the task description, the second part of task 1 consisted of establishing further search criteria and selecting one, two or several companies for comparison. The second part of task 1 can be seen in figure 30. All participants managed to determine the further requirements for the search (marked in yellow and seen on the left side of figure 30) without experiencing any problems. All participants noted that the system and the assistive technology clearly stated what had been selected and there was no need to continue with the previously mentioned *guessing game.*

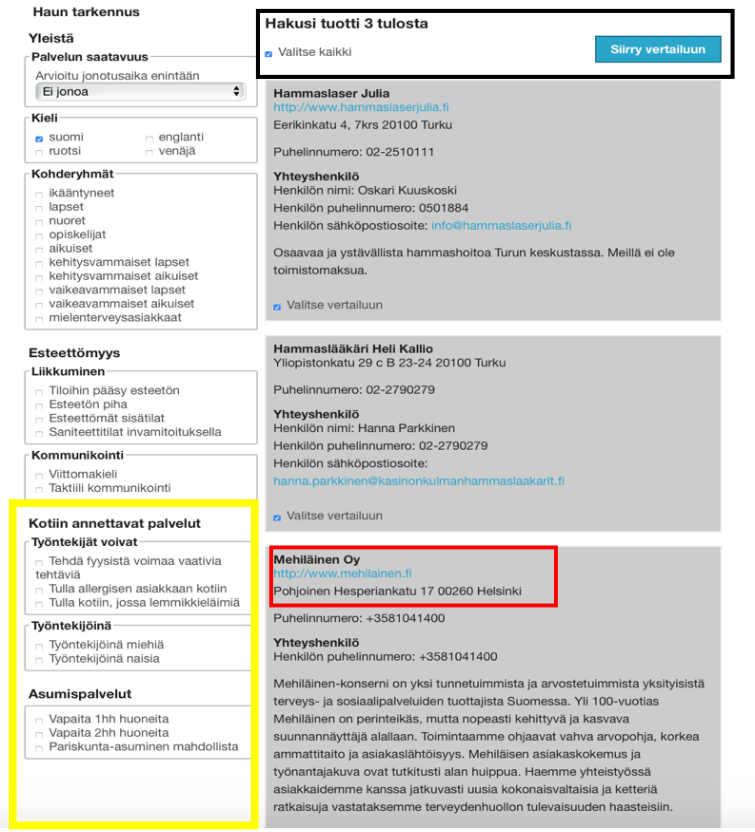


Figure 30: Part two of Task 1

Although none of the participants experienced any problems, there was still a modest impression of uncertainty among the participants. Consequently, as the task was to select a dental service for tooth removal, two thirds of the participants were confused as the further criteria concerning home services occurred. This obstacle did, however, not cause any major problems, since the majority of the participants were aware of the fact that the task concerned a service provided outside the home. Admittedly, the participants noted that this section could be left out in order not to confuse the user, as it is time consuming and can easily distract unexperienced users of the system.

In addition, the second part of task 1 caused uncertainty among the participants, as the task was to find a service provider from the city of Turku, participants were confused when the search results indicated that a desired service provider was located in Helsinki, illustrated in figure 30 with a red square. This resulted in confusion rather than in usability or accessibility problems, as participants were

aware of the fact that the company provided services in Turku. However, several participants noted that this might cause significant problems for people who are unaware of a company's location or address.

The second part of task 1 further holds problematic aspects that participants experienced as usability related issues. The usability related issue is illustrated in the top right corner in figure 30. Two thirds of the participants noted that the "*siirry vertailuun"* link which allows the user to proceed to the comparison interface, is neither informative enough nor well located. According to the participants, it was to some extent difficult to determine the number of companies that had been selected for the comparison. Several participants pointed out that in order to make the system more user friendly, the system should indicate how many companies the user has selected to proceed with. At present state, participants are experiencing problems in determining and remembering how many companies have actually been selected. In order to know the exact number, the participant must navigate through all the results in order to answer that question. However, it important to note that even if the system does not indicate the number of chosen companies, the system will not allow the user to proceed without having chosen something. According to a more experienced user, as the system generally follows an *up to down* structure, it is confusing and rather non-user friendly that the link is located above the companies and their corresponding information. Since the link is now located on the top of the site and the search results beneath, users must navigate back and forth in order to proceed to the next interface. This could in fact cause severe problems among people using screen readers, since they might not know or understand to navigate their way back to the top of the site (if they did not notice the link before selecting the desired companies). As a result, a number of participants were unaware of the location of the link, which resulted in frustration and unnecessary time consumption.

The third part of task 1 was to find and select a company or company representative who had received the highest rating. Additionally, the participants were asked to point out if and how they would be in contact with a preferred

company representative. The third part of task 1 is illustrated in figure 31. Even though the link *"siirry vertailuun"* caused confusion, it did not cause problems in proceeding to the final phase of the task. However, one third of the participants were confused as the table now consisted of nine rows, even if only three companies had been selected for the comparison.

Most of the participants experienced similar problems with the last part of the task. All participants experienced difficulties in reading and understanding the table, since the assistive technology first read the titles of the columns, followed by reading the rows one by one. The table seen in figure 31, eventually follows the same structure of the previously mentioned *guessing game,* where participants have to remember the exact order of the column titles in order to know what a specific information within the rows is referring to. All participants pointed out that the table was difficult to understand and hardly accessible. One participant using both a braille keyboard and screen reader, failed to access the table and had to view the information from the printable version. This significant problem may, however, be related to the assistive technology that the participant used, since after several attempts the participant managed the view the table. On the other hand, a more experienced participant pointed out that the titles of the columns were inadequately programmed, as the screen reader read the default label *"activate to sort column ascending".* According to the participant, this type of labeling is unacceptable and a commonly occurring issue, which results in problems for people who are using screen readers with a selected language other than English. The participant further points out that people using assistive technology,

particularly screen readers, tend to identify these “hidden problems in the code” which are not otherwise visual.

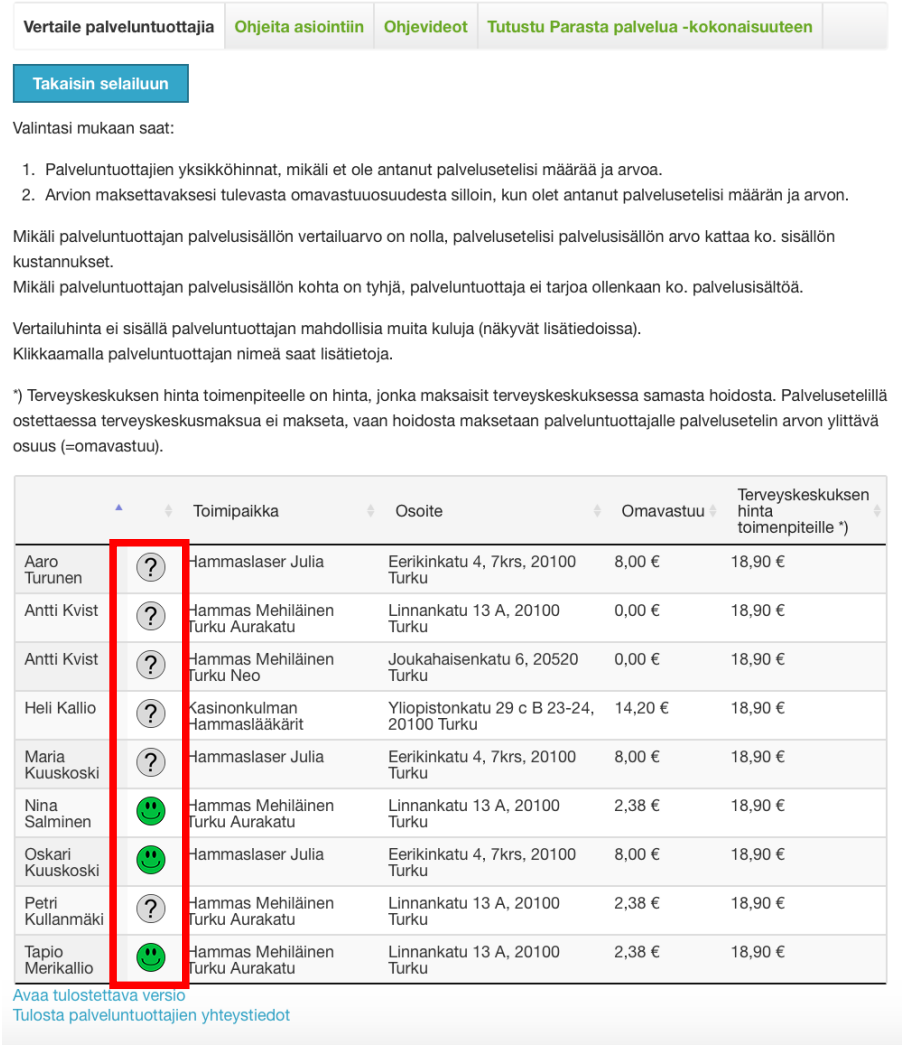


Figure 31: Part three of Task 1

As previously mentioned, the task was to select a company representative with the highest rating and to find the corresponding contact information. The rating, also known as the quality index, is illustrated in the second column of figure 31 and highlighted with a red square. All eight participants had difficulties in understanding the exact meaning or rating of a certain company representative, as the assistive technology only read the label as *not known, good* or *excellent,* depending on the color of the image. Additionally, one participant pointed out that since the table was not informative enough, it causes significant hindrances and would eventually

leave the user totally clueless without the task description. Nonetheless, participants were aware of the fact that the image held a link that could be interacted with, however, as the description of the image was limited, half of the participants decided not to interact with it. However, the other half interacted with the image by pure accident without knowing what would eventually happen. Consequently, none of the participants, who accessed the link of the image, could indicate the content of the new window that had opened on the screen, since neither the braille keyboard nor the screen reader could read it. A few participants could, however, tell that something had occurred since the image contained a link. A more experienced participant assumed that a new window had opened, while the others kept interacting with the table. The new window regarding the rating of a company representative can be seen in figure 32.

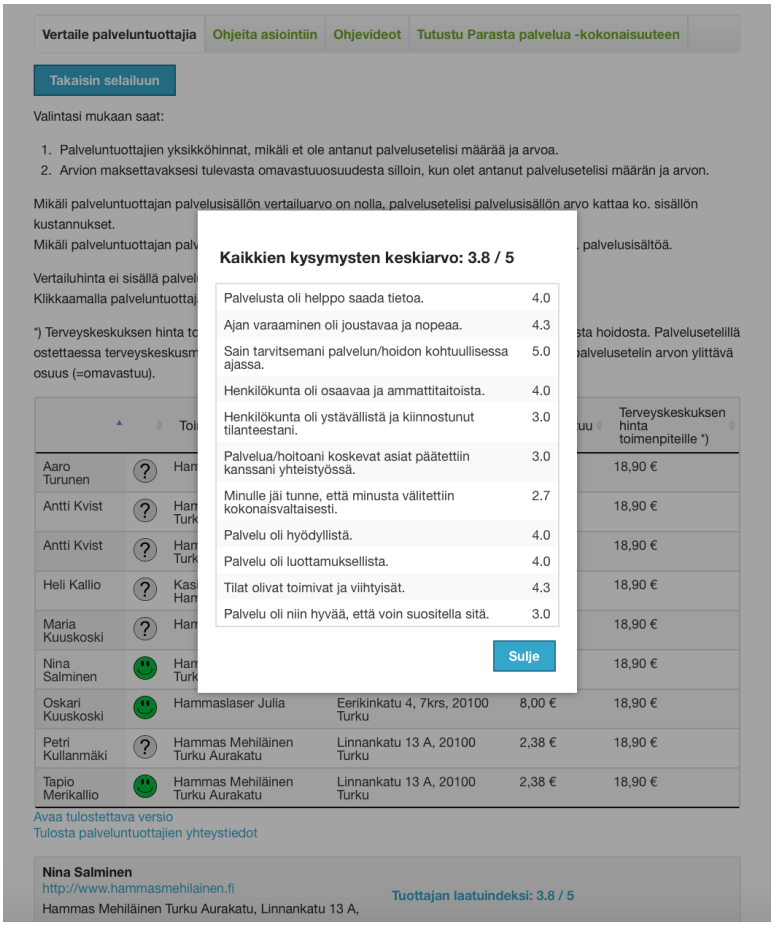


Figure 32: Rating of a company representative

The final step in task 1 was to find the contact information of a company representative with the highest ratings. All participants could, however, indicate that a company representative who had received a rating of *excellent,* was indeed the right one. Problems occurred when participants were asked to find the contact information, as two thirds of the participants were unable to access the new table that held further information about the representative. Although instructions are provided (see figure 31) and indicate that further information is provided when the name of a company representative is pressed on, all participants assumed that the system would be self-explanatory and therefore ignored the instructions. On the one hand, only one third of the participants were able to access and open the new table where further information was provided. However, this was done by pure accident after interacting with the link of the image and without following the instructions provided in the beginning. On the other hand, none of the few participants who managed to access the new table knew that a new table had been created below.

In conclusion, all participants experienced significant accessibility and usability related problems with task 1. Most of the identified problems are related to inadequate labeling of titles and followed by deficient placement of boxes and text fields. These problems caused both frustration and accessibility related hindrances. All participants noted that the system follows an inconsistent structure and often leaves the user clueless of what is expected to be done. Additionally, two thirds of the participants pointed out that the whole service price comparison section is illogical and requires the user to remember certain aspects in order to be complete a desired action. In other words, the system is not self-explanatory and leaves the user in a hopeless state, as the user will have to guess what is going to happen next. Similarly to people without sight impairment, blind people tend to ignore provided instructions and navigate their way through a site with the assumption that the site will be self-explanatory. Furthermore, a more experienced participant pointed out that "*a web service that should be accessible for disabled people, especially for people with sight impairment, should be programmed according to their mindset*

and logic, not according to the programmers. The end user should always be the center piece".

6.1.2 Findings from Task 2

Task 2 was about evaluating a service that a participant has received. However, this task could only be performed by half of the participants (four participants), since the task requires the receipt of a service voucher through the PSOP system. Participants were asked to search for a service provider and a service they recently received and complete the evaluation form, consisting of 11 predefined questions.

Task 2 is considered to be less challenging than to task 1. Undoubtedly, problems occurred among the four participants who were able to perform the task. After reading the task description for the participants, two participants were eager to address that they had never evaluated a service before and did not know it was a possibility. Additionally, none of the four participants experienced problems in accessing the link that leads to the evaluation form, seen in figure 33.

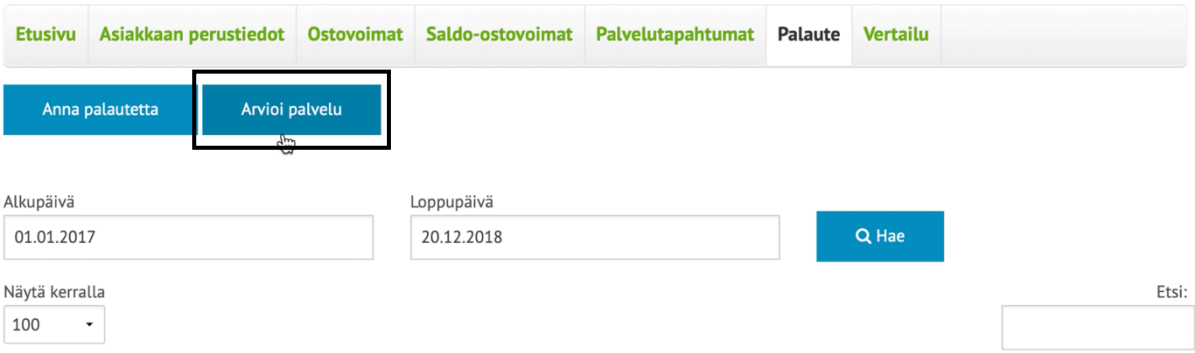


Figure 33: Proceeding to the evaluation

All four participants managed to find both the service provider and a service that they had recently received, without confronting any problems. However, severe problems began to occur immediately when participants started interacting with the evaluation form, seen in figure 34. All four participants noted that the titles of

the checkboxes were inadequately implemented on the form. This problem resulted in the previously mentioned *guessing game* and participants were required to remember and count the exact order of the titles in order to tick the proper checkbox. In other words, the assistive technology read the titles one by one, followed by reading the empty and uninformative checkboxes. Consequently, none of the participants could indicate if the right checkbox had been selected. The same problem followed throughout the whole evaluation form. The problem area of task 2 is illustrated as red square in figure 34. This resulted in frustration and eventually left the participants clueless.

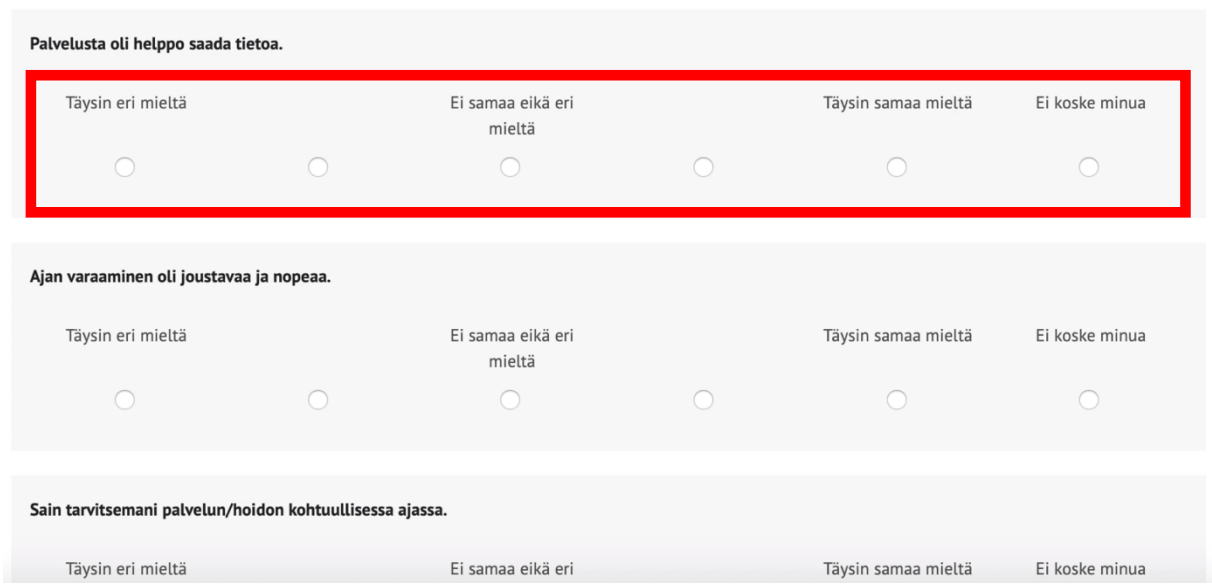


Figure 34: First three questions of the evaluation form

Additionally, participants experienced further problems due to the fact that each question held four titles and six checkboxes. The problem resulted in the already discussed *guessing game,* as participants were forced to count the exact or simply guess which checkbox referred to which title. Although participants tried to keep count of the titles and checkboxes, none managed to answer the form in a desired manner.

In conclusion, all the participants experienced severe problems throughout the evaluation form. After completing the 11 questions, all participants were still

unaware if they had answered the evaluation form as they initially desired. However, all participants attempted to keep count of the titles and tick the correct checkboxes by navigating back and forth between the questions. Despite this action, none of the participants could indicate whether a question had been completed in a preferable manner. In contrast to all the problems and inconsistency within the evaluation form, participants could, however, easily indicate that the evaluation had been successfully sent, seen in figure 35.

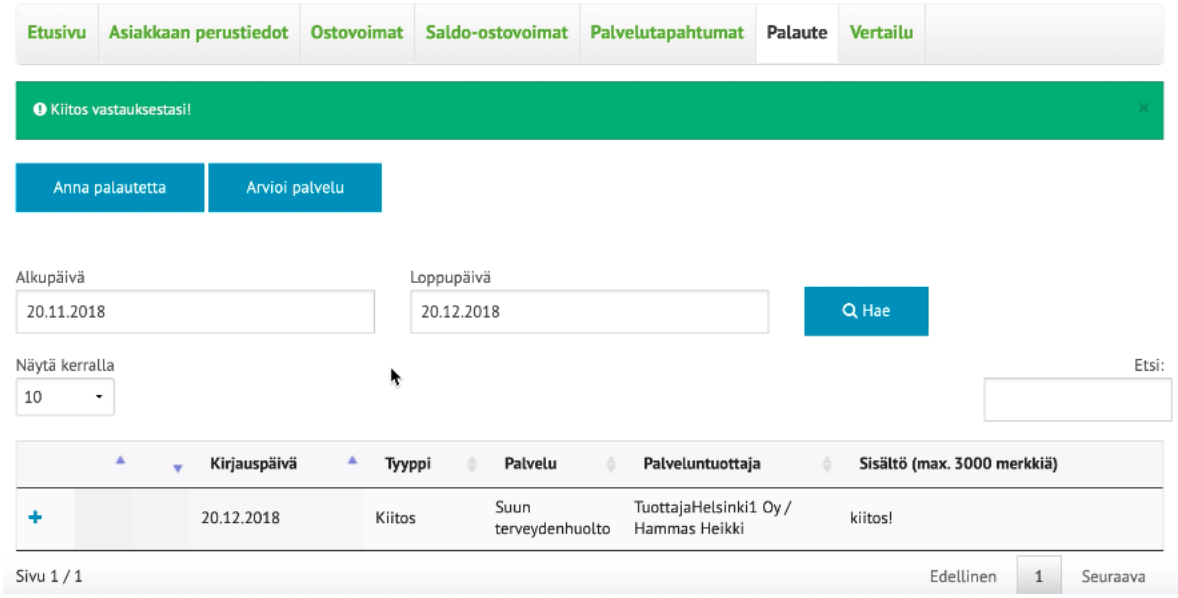


Figure 35: Successful submission of an evaluation

6.1.3 Findings from Task 3

Task 3 was about providing feedback to a recently received service. However, task 3 was similarly only performed by half of the participants (four participants), since the task requires the receipt of a service voucher through the PSOP system. According to the task description, participants were asked to provide feedback according to their own preferences, to a service of their liking. The content of the feedback was also left for the participants to decide. However, the only requirement was to ensure that the service provider would respond to the submitted feedback.

None of the four participants experienced problems in finding or accessing the feedback link, which can be seen in figure 36. However, one participant pointed out that they did not know or realize that submitting feedback was a possibility, otherwise the participant would have given a complaint to a previous service.

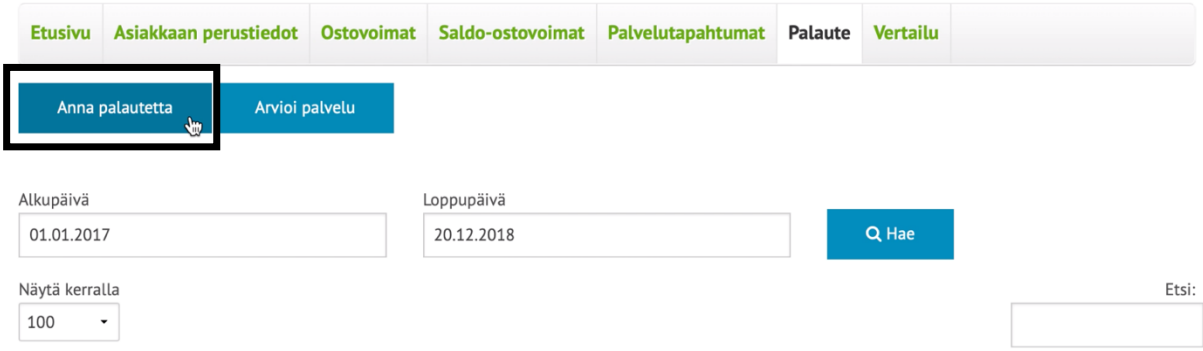


Figure 36: Interface of providing feedback

Firstly, none of the four participants experienced any problems, in selecting a service provider, service nor the subject of the feedback. Secondly, no problems occurred when participants wrote their preferred message in the field of contents. Thirdly, all participants were able to find the checkbox, which indicates that the service provider must reply to the feedback. Fourthly, all participants were able to successfully submit the feedback. The interface of the feedback form can be seen in figure 37, where the required criteria is marked with a red square.

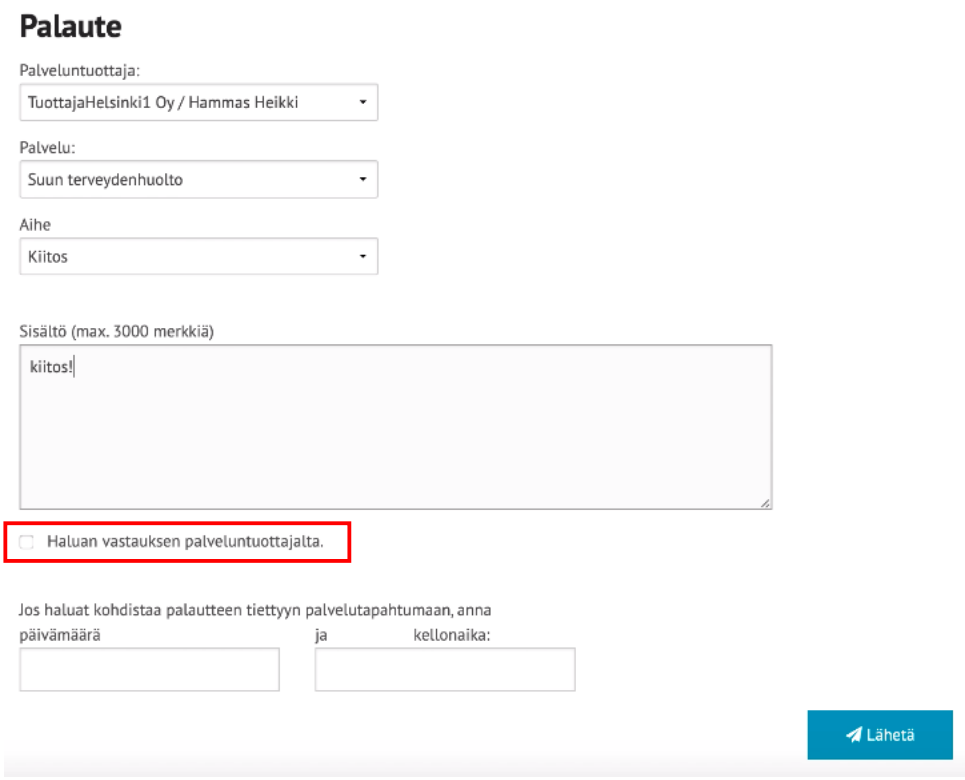


Figure 37: Providing feedback to a received service

Nonetheless, problems began to occur when participants were asked to tell whether the system now indicated that the service provider must provide an answer to the feedback. The request for an answer is illustrated as a red exclamation mark and seen in figure 38. However, none of the four participants could tell if the requirement had been satisfied as the row was not expanded. Furthermore, the assistive technology (both braille keyboard and screen reader) was not able to identify what the red exclamation mark referred to. Additionally, two participants pointed out that the column title of the exclamation mark was empty (without any information) and therefore the assistive technology could not indicate what the section meant or what the exclamation mark was referring to. Another user pointed out that the table, illustrated in figure 38, possessed the same problem seen in task 1, where the title of the columns contained the default labeling *"activate to sort column ascending".* These observed problems left participants confused as they were unsure if the only criteria had been satisfied.

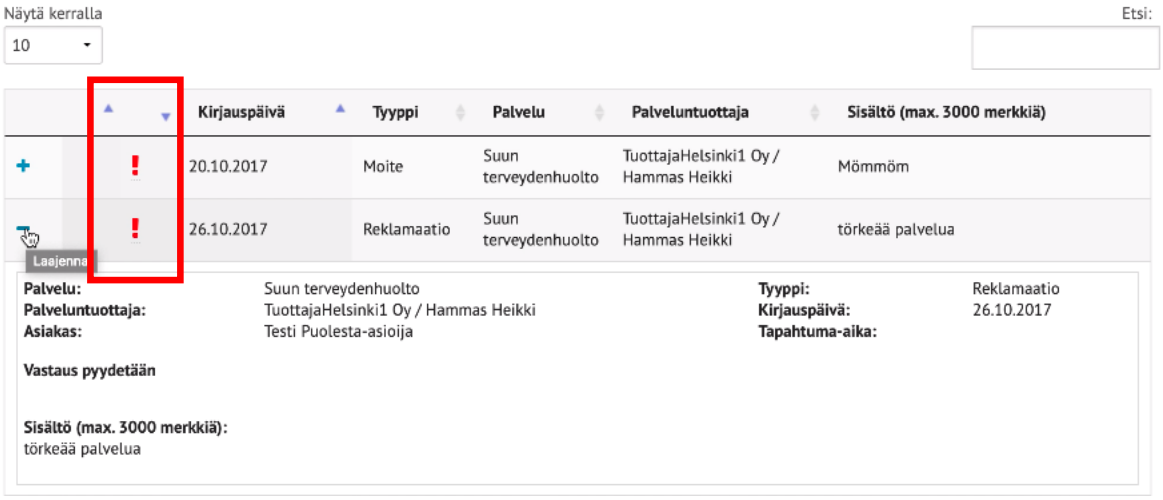


Figure 38: Feedback provided by the customer

The think-aloud method was similarly used in task 3 in order to gather additional data. Participants were asked if they could access a specific row of the table and further tell if all the information of the feedback had been stored properly. Half of the participants were able to expand the row without any major problems and identified that the information was indeed stored. Moreover, by expanding the row, these participants were now able to determine that the only criteria had been satisfied.

The remaining participants were not able expand the row and eventually decided to give up. However, after all the tasks were completed, the two participants were encouraged to return to the feedback section and further explore the interface. As a result, the two participants managed to expand the row and accessed the additional information of the feedback. This was, however, done by pure accident as the participants accidentally understood that the exclamation mark contained a link and could be interacted with. Despite accessing the additional information, none of the two participants understood how they had managed to expand the row. Nonetheless, the two participants were now able to determine that all information of the feedback had been stored and that the task was successfully completed.

In conclusion, none of the four participants experienced any problems submitting the feedback. Problems occurred when participants were asked to identify if the service provider had to reply to the submitted feedback. As a result, participants were unable to answer nor indicate whether the task had been completed when the rows were not expanded. However, half of the participants were able to access the expanded view, whereas the other half experienced severe problems and found it by pure accident. The participants noted that although the row was eventually expanded, uncertainty still remained as they did not know for sure what had happened when they interacted with the exclamation mark. In other words, the feedback section is rather frank, however, the table still possesses similar problems as seen in task 1 and eventually leaves the user clueless.

6.2 The difference between the findings on mobile and computer

As previously mentioned, the user test was conducted on both the mobile and the desktop platform. However, due to time constraints, task 1 was the only task conducted on both devices. Despite the time constraints, the user testing on the mobile device was conducted by all eight participants. Generally, people using assistive technology tend navigate through the whole site in order to get familiar with website and to know how the content is positioned. This is done in order to feel comfortable with the site, as it also reduces the risk of entering personal information in wrong places. This similarity can be seen among people without sight impairment. Undoubtedly, it can be clearly stated that participants experienced similar problems on both devices.

On the mobile platform, accessibility related problems started to occur as participants were asked to enter the information of the service voucher. Participants noted that the built-in screen reader on Apple devices, reads the section of the service price comparison in a similar manner as the screen reader on the desktop platform. Additionally, the section left participants clueless and confused if the right information had been entered. All participants noted that this

specific section of the system was inconsistent and did not follow the same structure as in the earlier section. As a consequence, half of the participants were unable to enter the quantity nor the value of the service voucher without guidance. Noticeably, participants conducting the first part of task 1, experienced similar problems on both platforms.

The second part of task 1 did not cause any problems for the mobile users. Participants were able to determine the preferred criteria and several participants pointed out that they could clearly indicate which checkboxes had been selected. Problems began to occur when participants were asked to select one, two or several service providers for the comparison. Participants had no problem finding the service providers, however, problems occurred as participants attempted to move to the comparison, as the participants were unable to find the link *"Siirry vertailuun"* that would take the user to the comparison. Participants pointed out that they had no idea where the link was located. On the contrary, half of the participants had noticed the link before browsing through the service providers but pointed out that this may cause severe problems for users who overlooked the link. In addition, two participants thought that the link would have been located after the listing of the service providers, as they assumed that the system would follow a *from top to bottom* structure. These participants further noted, that the system was not informative enough, as it did not indicate the selected number of service providers that was selected for the comparison. Noticeably, these above-mentioned problems are generally experienced both platforms.

The third part of task 1 caused problems from the very beginning. After proceeding to the comparison, one third of the participants found that the screen reader had placed them randomly in the middle of the site. This caused severe problems and left the participants distracted and confused. These participants had tremendous problems figuring out what the information on the site was referred to. However, after using the think-aloud method, all participants were able to understand the content. Noticeably, the table in figure 31 caused major problems on both platforms. All participants pointed out that the column titles were inadequately

programmed on both platforms, resulting in the previously discussed *guessing game* where participants experienced major problems in accessing the information of the table. Consequently, two thirds of the participants noticed that the information on the rows indicated which column number it referred to. However, as the screen reader only indicated the number of the column and not the title, participants remained confused. As an example, none of the participants were able to indicate what the difference between the amounts shown on the rows.

Similar accessibility related problems were identified on both devices when participants were asked to identify the rating of a company representative. Half of the participants noticed that the second column held images, however, as the screen reader only indicated them as *not known, good* or *excellent,* participants were confused and unsure whether to interact with it. The other half of the participants completely ignored the image link and were therefore unable to identify the rating of a company representative. However, participants who decided to interact with the image link, were unaware of the fact that a new window had opened on the screen, as the screen reader did not notify it. Some of the participants, however, assumed that something had happened but were unable to identify it. The similarity of the problem can be seen on both platforms.

Additionally, the third part of task 1 caused severe problems among all eight participants, as visually impairment people tend to navigate their way through a site relatively fast, disregarding instructions and expecting the website to be self-explanatory. Two thirds of the participants noted that the instructions in the beginning of the site clearly states that "*by interacting with the name of a company representatives, further information will be provided",* none of the participants could access the information by interacting with the representative's name. Participants pointed out that this is clearly an accessibility related problem, since the section does not contain a link to interact with. However, participants who had earlier decided to interact with the image, noticed that a new table with further information of the company representative, had been created under the original

table. Despite accessing the table, participants pointed out that it was done by pure accident

In conclusion, similar problems can be identified on both devices. Firstly, entering the information of the service voucher causes severe problems on both devices, since the previously mentioned *guessing game* is present. Secondly, participants are left clueless of how many companies has been selected for the comparison. Thirdly, participants are experiencing severe problems in accessing and understanding the information provided in the table of the company representatives. Additionally, none of the participants could access the further information of a company representative according to the instructions provided by the system. Consequently, several aspects of the system are found by pure accident, without knowledge or without the screen reader informing about it.

6.3 Findings from the SUS questionnaire

As previously noted, the predefined System Usability Scale (SUS) questionnaire was supposed to be sent to the customers of the PSOP system. However, due to the General Data Protection Regulation, the questionnaire was not distributed. As a result, the SUS questionnaire could, however, be used as a post-test questionnaire, as it was read aloud to the participants after the test session. Generally, it requires a large sample in order to gather reliable data from the SUS questionnaire but according to Brooke (2013), it has been demonstrated that the SUS questionnaire is in fact reliable even with a small-scale sample. Therefore, even with a small sample of eight to twelve participants, the SUS questionnaire is able to measure the perceived usability of a tested system (Brooke, 2013).

The SUS questionnaire consisted of ten predefined questions regarding the systems usability, ranging from strongly disagree to strongly agree. The distribution of the answers among the eight participants are illustrated in table 3. As seen in table 3, the perception regarding the usability of the PSOP system varies significantly. On

the one hand, approximately 75 percent of the participants perceived the system as unnecessarily hard to use, whereas the remaining 25 percent either disagreed or gave a modest answer to the claim. On the other hand, 62,5 percent of the participants experienced the system as exceedingly hard to use, while 37,5 percent disagreed or strongly disagreed. As a result, 56,25 percent of the participants experienced that the system was difficult to use, as the remaining 43,75 percent either disagreed or had a modest perception of the systems usability. However, none of the disagreeing participants, indicated that the system would have been relatively easy to use. Consequently, the majority of the answers leaned towards a more negative perception of the systems usability. Additionally, 62,5 percent of the participants noted that they would probably need technical assistants in order to use the system. Similarly, 50 percent of the participants noted that they would not prefer to interact with the system frequently, whereas the remaining 50 percent of the participants had a modest answer to the claim.

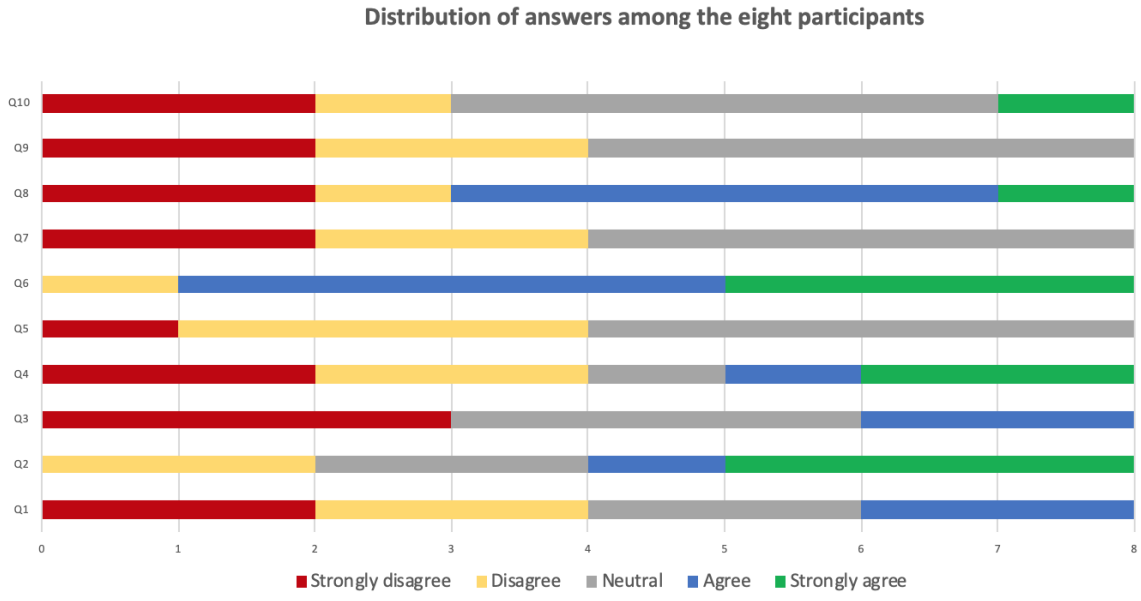


Table 3: Distribution of answers among the eight participants

As observed in all the tasks, especially in task 1, participants experienced that the system was highly inconsistent and 87,5 percent of the participants either agreed or strongly agreed to this claim. As a result, the SUS questionnaire strongly supports the findings regarding the systems inconsistency, which was identified during the

test sessions. Undoubtedly, all the participants felt particularly unconfident when using the system and believed that most people would experience difficulties learning how to interact with the system. In addition, 50 percent of the participants either strongly disagreed or disagreed to the claim whether the different areas and sections of the system worked well together, while the remaining 50 percent, however, gave a modest answer to the claim.

In conclusion, the score per participants, as well as the average score from the whole test session, is illustrated in table 4. As seen in table 4, the SUS score between the eight participants varies slightly. The variance in the score might be due to the previously mentioned expertise in IT among the participants. However, since the participants were allowed to conduct the test on their own devices, the results are reliable. Table 4 illustrates the scale of the SUS score on the y-axis and the eight participants on the x-axis. It can be clearly stated that participants p5, p7 and p8 were able to conduct the test without any major hindrances and therefore received a higher SUS score. On the other hand, participant p1, p2, p3, p4 and p6, evidently experienced major problems. The red line in table 4, illustrates the average SUS score that was calculated from the questionnaire. The average SUS score, that resembles the usability of the PSOP system is 38,8. As mentioned before, the scoring system for SUS ranges from 0 to 100 and according Brooke (2013) the average SUS score for a system is generally around 70. As a result, the score of 38,8 is considered to be significantly below average and to resemble a poor usability (see figure 27). Consequently, 62,5 percent of the participants received a score below 50, which indicates that there are indeed usability related problems within the system.

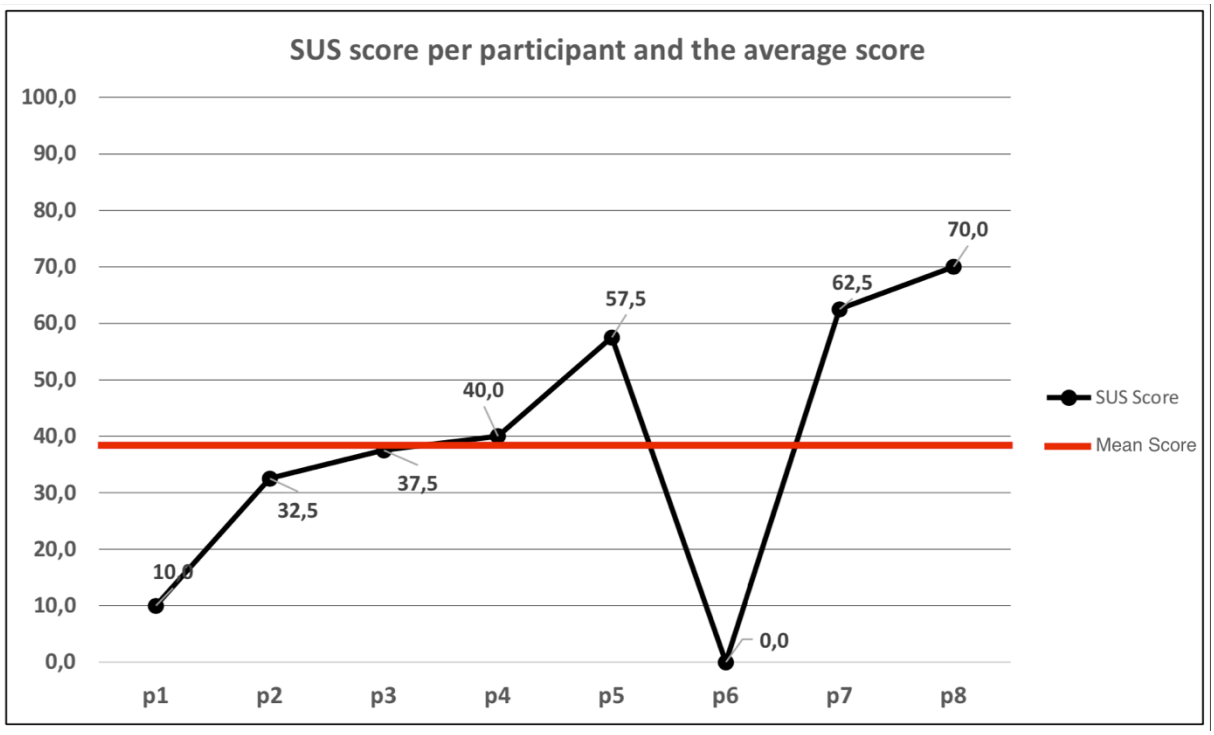


Table 4: SUS score per participant and the average score

6.4 Evaluating the problems

The evaluation of the identified problems was analyzed by the severity, perception and the number of times a problem occurred. The severity of the encountered problems was evaluated by analyzing how much the problem affected the usability and the accessibility as well as how the participants managed to prevail the problem. The perception of the problem was analyzed in order to identify how the participants perceived a problem. The number of encountered problems was identified during and after the test session by using the think-aloud method and by analyzing the recorded material. Analyzing the recorded material enabled for identifying how frequently a problem occurred.

Firstly, it can be clearly stated that all participants experienced problems with the inadequately positioned titles, following by the non-informative text fields and boxes. Noticeably, the severity regarding this problem is significantly high as all participants experienced it in several areas throughout the systems interface. Additionally, this specific problem resulted in the frequently mentioned *guessing game* that left the participants clueless and frustrated. Secondly, accessing the

information in the tables in various sections of the system, caused major problems. The prominent severity lies in the image links, as none of the eight participants could identify the quality index of a company representative, nor access the newly opened popup-window. Additionally, the mobile test session identified the fact that the system indicated that further information can be accessed by a certain action, the assistive technology was, however, not able to access it by following the provided instructions. Thirdly, all participants experienced problems accessing the systems various tables. This caused severe problems and also resulted in the previously mentioned *guessing game*. Additionally, the majority of the participants considered the system to be inconsistent, which according to the participants, is a major usability problem for people who are using assistive technology, such as screen readers.

In conclusion, the perception of the encountered problems among the participants varied slightly. However, none of the participants were able to perform task 1 and task 2 with a success rate of 100 percent. With the use of the think-aloud method, participants were encouraged to further make an effort to overcome certain problems. The think-aloud method turned out to be an efficient tool in identifying underlying problems. Undoubtedly, the most error-prone sections of the system are identified in task 1 and task 2, where the inadequate labeling and accessibility of images caused the majority of the problems. The least error-prone section of the system is identified in task 3. However, immediately as the participant accesses the table to review further information, major problems occur. This similarity can be seen in all the tables within the system.

7. Discussion

The purpose of the research was to identify both accessibility and usability related problems within the PSOP system, among disabled people. The aim was to conduct a user test, involving real users, in order to accomplish the goal of identifying the real problems that the systems users face.
The following chapter provides answers to the two research questions that were presented in the introduction chapter. Firstly, chapter 7.1 presents answers to the first research question, where a summarized overview of the problems is presented. Secondly, chapter 7.2 presents answers to the second research question, including reasoning behind the identified problems and further analysis of required future actions.

7.1 Research question RQ1

RQ1: *Although a technical test regarding the WCAG has previously been conducted, what kind of problems do users with sight impairment experience when interacting with the PSOP system?*

Despite the fact that a technical study has earlier been conducted, both accessibility and usability related problems were identified by conducting a user testing with eight blind participants. The technical study analyzed problems such as visual appearance, navigation and full accessibility with keyboard functions as well as attempted to present the content in a simplified format. However, after successfully conducting a user testing with visually impaired people, it can be well-argued that significant problems still exist, as the perception of usability and accessibility is divergent. Consequently, a large number of various problems were identified during the test as well as after by analyzing the recorded material. The majority of the participants identified the same problems within each task, which furthermore confirms the theory behind Nielsen's perception of involving a small sample of users for user testing (Nielsen, 2000). Additionally, inconsistency related

problems were identified. According to the findings, the PSOP system does not follow a consistent structure throughout the system, which is vital for blind people using assistive technology. The problem regarding inconsistency is especially visible in task 1. The findings further specify that blind people tend to memorize the structure of a website and assume that the content follows the same structure. However, the current stage of the system does not follow a strict structure, which is clearly a usability related problem. Furthermore, based on the findings, the system is missing a much desired, self-explanatory content, which would ease the navigation and as well as the overall efficiency. Nevertheless, the system partly holds clear instructions that informs the user how to interact with the system. However, similarly to non-disabled people, visually impaired people tend to ignore these instructions and expect that the system will indicate what actions are required from the user. The self-explanatory problem is especially prominent in task 1 but can also be seen in other tasks, where these instructions are not provided. Consequently, the self-explanatory problem will probably occur in the purchasing power and in the balance of the purchasing power sections (presented in chapter 3.1), which were not included in this study due to limitation reasons.
Nonetheless, the most severe and error-prone problems were identified in the labeling and programming of titles and text boxes respective fields. According to the findings, the PSOP system holds an inadequate programming of titles and fields, which are inefficiently placed on the site and hardly identified by assistive technology. This problem is clearly an accessibility related problem; thus, it also causes significant usability related problems among the users of the system. Another accessibility related problem is the access to specific content on the system. The user testing identifies that tables, images and further information are nearly inaccessible. According to the results, especially images and interaction with various tables, causes significant accessibility problems. Consequently, visually impaired people using assistive technology such as braille keyboards or screen readers, are unable to interact or observe the information behind images and image links.

In conclusion, all the identified problems, are a reason why participants experience the PSOP system as non-user friendly and inconvenient to use. Undoubtedly, as all these problems are present, the system eventually creates an atmosphere that forces the user to guess or assume what the system expects the user to do. Additionally, as several participants pointed out, the system leaves the user clueless and a so-called *guessing game* occurs. Nevertheless, the findings from the System Usability Scale questionnaire, further supports the non-user friendly perception of the PSOP system, as the participants rated the usability of the system with an average score of 38,8, from a maximum score of 100. However, the perception of the systems usability varies depending on the users, which can be due to their IT experience or former knowledge and perception of the system.

7.2 Research question RQ2

RQ2: *What are the reasons behind the identified problems within the PSOP system and what kind of actions should be done in order to eliminate the problems?*

The answering of the second research question is important to address, since it is essential to consider what caused the problems to occur. All the identified problems have relatively clear reasons why they might have occurred. However, it is relevant to also consider that there might be underlying factors that are eventually unknown. Firstly, the problem regarding inconsistency is a clear usability related problem than may not in fact be a problem for the general public. However, since people with disabilities and especially blind people, are constrained to navigate their way through an interface in a chronological order, it can easily cause significant usability related problems. The reason behind this observed problem might be that the system is in first hand designed for the general public and therefore not designed to be efficiently used by assistive technology. Secondly, the problem regarding the systems efficiency and its uninformative aspects, is considered to be a usability related problem. Although the system holds instructions, users tend to ignore them. According to the results, the system is not informative enough and does not clearly indicate or express what action the user is

expected to do. The reason behind this problem is significantly clear. The system has insufficient labeling of different titles, text fields, text boxes, links and images. In other words, visually impaired people are experiencing severe problems in knowing and understanding what actions are supposed to be done. Thirdly, the problem concerning the positioning of titles and different text fields and text boxes, can be considered as an accessibility related problem. The reason behind the problem is relatively clear. The programming and the positioning of titles and different fields, in various sections of the system, are inadequately defined and therefore causes severe problems. Nonetheless, it can be assumed that this problem is further caused by the fact that the system is initially designed for the general public. This specific problem can be witnessed throughout the system, as it is particularly visible in task 1 and task 2. Fourthly, the problem of accessing tables and further information in images, can be seen as an accessibility related problem. The reason behind the problem of accessing information provided in tables is generally difficult to determine and there might be underlying reasons for its occurrence. The problem might partially be due to the assistive technology that was used, since the braille keyboard was unable to access the information. However, the reason behind the severity of the problem lies in the fact that column names were programmed in the titles and not in the columns and rows itself. Another reason behind the problem is that the screen reader did not indicate which title a specific section in the table was referred to. Instead, the screen reader only read the number of the column without any additional information. Additionally, the default labeling of columns, is another problem that is caused by inadequate programming and commonly a hidden problem that screen readers are able to identify. Furthermore, the problem of accessing images and the information behind the image link, has a remarkably clear reason. Since the information of an image is provided in a new popup-window that appears on the screen, neither the braille keyboard nor the screen reader can access or detect it. This problem is clearly caused due to inadequate design and programming of the system.

As previously mentioned, it is relevant to consider that there might be underlying and non-observed factors that may have caused some of the identified problems.

Problems might have occurred due to the participants IT knowledge, former knowledge of similar systems or other experiences or expertise that the participant holds. It is also important to address that problems might have occurred because of underlying user circumstances, anxiety or due to the fact that some participants took part in their first user test session. However, factors such as test environment and recognition of assistive technology among the participants, can be clearly excluded from causing the problems. This aspect was one of the corner stones of the study, as the location of the test sessions and use of preferred assistive technology was left for the participant to decide.

Moreover, in order to establish efficient usability and foremost to ensure that the system is accessible, certain actions are needed to be done. To establish efficient usability, the system will have to follow a consistent structure throughout the whole interface. By designing a *from up to down* structure, people using assistive technology will experience a more user-friendly interaction with the system. At current state, the PSOP system has a partly consistent structure but according to participants, the structure can suddenly change completely, causing significant problems. Additionally, the reprogramming and repositioning of titles, different fields and boxes would tremendously increase both the usability and accessibility of the system. By addressing these severe problems, the system could effortlessly become more self-explanatory and particular instructions could eventually be removed. Moreover, the reprogramming and renaming of tables and image links, is another aspect to address in order to ensure both usability and accessibility. Firstly, by renaming the specific sections in a table with a title, instead of the column number, would increase the usability significantly. Secondly, the system should be redesigned in order to ensure that the assistive technology understands that something can be interacted with. At current state, these instructions are present but assistive technology such as screen readers are unable to receive the information, although instructions would have been followed. Thirdly, by redesigning the images and image links, in a way that the assistive technology would access it, is an essential aspect to consider.

According to the findings, both the usability and accessibility can be increased and ensured if the system is reprogrammed and redesigned according to the mindset and logic of its whole customer base, including people with disabilities.

In conclusion, the PSOP system is fairly usable, accessible and satisfies the required guidelines provided by WCAG, the study identifies that by conducting a user testing with real users, additional problems can be distinguished. It can clearly be noted that, although the PSOP system is considered to be *accessible on paper*, user testing is able to identify problems that real users face when interacting with the system. User testing can also identify that people with disabilities have a diverse perception regarding both the usability and accessibility of the system.

8. Conclusion

The aim of this study was to identify both usability and accessibility related problems within the PSOP system, among visually impaired. Additionally, the importance of the study arises from the fact that Valtiovarainministeriö (2018) has determined specific rules and Web accessibility guidelines for content provided by the public sector. Therefore, it is important for Kuntien Tiera Oy to take the observed problems into consideration, as the PSOP system undergoes further development after the second quarter of the year.

Both usability and accessibility related problems were identified using a mixed method of qualitative and quantitative techniques. Usability problems were identified by commonly used techniques and by combining methods such as user testing and the think-aloud method. However, as prior research points out, the same methods can also be used in order to determine accessibility related problems (Brajnik, 2008). According to Brajnik (2008), usability methods are in fact a valuable set of techniques that can be used to address both usability and accessibility problems. Consequently, the tools and techniques used in this study resulted in distinguishing both usability and accessibility problems.

Another goal of the study was to determine whether problems were still present, although a technical accessibility study had already been conducted. According to prior research, accessibility related problems are clearly identified when real users are involved, therefore a user testing involving real users was conducted. Consequently, an automated test is able to address the problems in the code and determine how accessible the system is on paper, however, this does not necessarily indicate that the system would be flawless and accessible in reality. According to Cheng and Mustafa (2014), user testing is still a method that is used more than half of the times when it comes to detecting problems. Undoubtedly, the importance of addressing both usability and accessibility related problems, lies in the test approach (Cheng & Mustafa, 2014).

This study has proven the fact that in order to identify the real usability and accessibility rate of a system, real users, preferably users from the target group, should always be involved and tested.

8.1 Limitations

The study involved limitations throughout the process. Limitations such as recruitment process, participants, arrangements and the used methods were present. The whole recruitment process changed drastically throughout the process, as the initial plan was to receive participants from Kuntien Tiera Oy, mailing lists and from the City of Turku. However, due to certain difficulties, alternative solutions had to be done. Another limitation in the study was the number of participants. The initial plan was to involve between 10 and 15 participants in order to gather reliable data. It was also intended to involve participants, who were all users of the PSOP system. However, as Nielsen (2000) points out, that only five participants are enough, the limitation did not interfere with the results. Nonetheless, by having a larger sample size, with all participants being users of the system, would have further increased the reliability of the results. Another limitation was the arrangements for the study. The study was evaluated and moderated by a single person. As a result, the observed problems as well as the results might be affected by the persons knowledge and experience, as the study may have neglected certain problems. Additionally, another limitation was the methods used in the study. It can be clearly notified that user testing turned out to be the most efficient method in order to identify the real usability and accessibility related problems. However, methods such as the Heuristic Evaluation and Cognitive Walkthrough were not applied, as they were not identified as suitable methods in order to identify problems, since neither of the methods require real users. In addition, the findings from the SUS questionnaire might be biased, as the SUS is not strictly intended to measure usability among people with disabilities. However, there are limited research and evidence indicating whether it cannot be

used. Therefore, questions concerning the reliability of SUS and whether it is a suitable method in measuring usability among visually impaired people arises.

In contrast to the limitations, some strengths were also present. The most noticeable strength of the study was that the test sessions were conducted according to the participants requirements. Therefore, the test environment as well as the results were closely representing a realistic interaction with the system.

8.2 Recommendations for Kuntien Tiera Oy

The findings from this study provide an understanding on how visually impaired users utilize different assistive technology, such as braille keyboards and screen readers in order to interact with the PSOP system. The study identifies various problems as well as how the assistive technology works throughout the system. Visually impaired users are heavily dependent on accessible web design and are expecting the system to be both self-explanatory and consistent, in order to utilize the system to its fullest. The study discovered several aspects within the PSOP system that were perceived as difficult or as exceedingly challenging by the participants. Common challenges between the participants were identified throughout the study. Despite the fact that not all of the tasks could be completed by all participants, the findings can, however, determine that usability and accessibility related problems are present. However, the difference between the participants IT knowledge might be a factor that needs to be accounted for.

The study proposes future recommendations for the PSOP system, as the system is undergoing further development after the second quarter of the year. Kuntien Tiera Oy should take the findings from this study into consideration and attempt to address the identified problems. By eliminating the problems found in this study, both the usability as well as the accessibility would increase significantly.

An essential method that is recommended to follow is the so-called A/B testing, also known as split testing, which has been used by a wide range of online services

(Gui, Xu, Bhasin, & Han, May 18, 2015). The fundamental idea of A/B testing is to compare two versions of the same system in order to address problems and to identify which elements of the system are important and which needs to further undergo certain procedures (Gui et al., May 18, 2015). A/B testing is a valid method for Kuntien Tiera Oy to consider, especially as the same problem of accessing tables may occur in other sections of the system as well (see figure 16). Additionally, the problem regarding the systems lack of being self-explanatory is also likely to appear in other areas such as in the purchasing powers section, see figure 14.

After addressing the observed problems, a good alternative for Kuntien Tiera Oy would be to run an A/B test and compare the two versions with each other in order to identify whether the problems are eliminated or not. Additionally, an automated test is also recommended, for the purpose of ensuring that accessibility related problems are satisfied according to WCAG and the *Accessibility Directive*. However, as noticed in this study, the automated test should always be followed by user testing, in order to analyze whether the system is in fact accessible.

In conclusion, a similar study could be conducted with a larger group of participants in order to receive broader results. It would also be recommended to further test the system on different mobile devices, as several participants pointed out that it will eventually be the preferred device in the future.

8.3 Recommendations for future research

There are several recommendations for future research, as the demand increases to ensure accessibility for websites provided by the public sector. Firstly, although Nielsen (2000) points out that only five participants are needed to gain reliable data, a study involving a significantly larger sample would contribute with interesting results. Secondly, by conducting a study involving novice and experienced users, would result in an interesting comparison of how users perceive both the usability and accessibility of a certain system. Thirdly, by further extending the methods used in this study, e.g. with additional evaluators and moderators, the

possibility of gathering additional data increases. Nonetheless, in order to ensure both usability and accessibility, a test session involving real users is still highly recommended.

Consequently, the variety of accessibility methods could be further discussed and improved. Although there are methods specifically intended for accessibility testing, the study shows evidence that other methods, such as usability methods, can be utilized in order to identify accessibility related problems. However, it is acknowledged in several sections throughout the study, that accessibility methods do not guarantee that a system or website becomes accessible to the whole customer base. Therefore, future research should try to identify and establish universal accessibility testing methods to ensure accessibility by everyone, regardless of disability.

9. Svensk sammanfattning

1. Inledning

Betydelsen av användbarhetstest på webben har ökat avsevärt i över ett decennium och testen förblir en central del av designprocessen, eftersom det tekniska genombrottet fortsättningsvis påverkar företag och organisationer. Intressenter och andra beslutsfattare har insett betydelsen av användbarheten och mängden pengar och tid som investeras i designprocessen, för att säkerställa användbarheten, har ökat med enorma steg. Detta är i synnerhet viktigt i och med att företag och organisationer försöker allt mer uppfylla kundens användarupplevelse.

Utöver användbarhet, har även efterfrågan på en tillgänglig webb fortsatt att öka. Tidigare forskning påpekar att tillgänglighet i likhet med användarbarhet, har under de senaste 10 åren fått mycket uppmärksamhet av både forskare och av andra intressenter. Företag och organisationer som skapar webbinnehåll, har gjort stora insatser för att skapa webbtillgänglighet för personer med olika funktionsnedsättningar, som till exempel synskadade. En stor del av webbinnehållet är trots detta fortfarande icke tillgängligt och framför allt till en del oanvändbart för personer med funktionsnedsättningar. Denna otillräcklighet i webbtillgängligheten kan högst antagligen bero på en ologisk designprocess, där hjälpteknik så som skärmläsare inte tagits i beaktande. Detta resulterar i att personer, som använder sig av diverse hjälpteknik, upptäcker problem med missledande innehåll, otillräcklig struktur och brister i bildbeskrivning. Dessa personer har i allmänhet svårt att få tillgång till tabeller och annat innehåll. Fastän förbättringar ständigt görs för att säkerställa webbtillgänglighet, så stöter dessa personer fortsättningsvis på problem med webben.

Utöver detta kommer betydelsen av att skapa en tillgänglig webb fortsättningsvis att öka eftersom ny teknik framträder på marknaden. Som en följd av detta har flera länder genom politik och lagstiftning bestämt riktlinjer för webbtillgänglighet. Finland har på liknande sätt följt Europaparlamentets direktiv för säkerställning av

tillgängligheten till webbplatser och mobila applikationer, som producerats av den offentliga sektorn. Tillgänglighetsdirektivet kommer gradvis att tillämpas av den offentliga sektorn från och med den 23 september 2019.

Denna avhandling görs på uppdrag av Kuntien Tiera Oy. Företaget producerar IKT-tjänster för att stöda utvecklingen av sina kunders dagliga verksamhet. Trots det stora utbudet av olika tjänster, fokuserar denna studie på systemet *Palveluseteli- ja ostopalvelujärjestelmä* (PSOP). I och med den nya lagstiftningen, bör Kuntien Tiera Oy och PSOP-systemet följa tillgänglighetsdirektivet (som följer riktlinjerna för WCAG), vilket företaget är medveten om.

Inom ramen för uppdraget från Kuntien Tiera Oy, önskar företaget att PSOP-systemets tillgänglighet och användbarhet skall testas av synskadade användare. Ett användartest utfördes för att analysera om PSOP-systemet uppfyller de riktlinjer som definierats för webbtjänster som tillhandahålls av den offentliga sektorn. Studiens tyngdpunkt ligger i att identifiera hur tillgänglig systemet är bland synskadade användare.

Nedan presenteras de forskningsfrågor, som kommer att besvaras:

1: *Fastän ett tekniskt test gällande WCAG tidigare har genomförts, vilka problem stöter synskadade användare på när de använder sig av PSOP-systemet?*

2: *Vilka är orsakerna till de identifierade problemen inom PSOP-systemet och vilka åtgärder bör vidtas för att de skall elimineras?*

2.1 Funktionsnedsättning

Det är ytterst svårt att definiera funktionsnedsättning under en allmän term. Däremot kan begreppet definieras som ett paraplybegrepp som består av olika nedsättningar, som exempelvis aktivitetsbegränsning. I sin enkelhet hänvisar funktionsnedsättning till ett problem i en persons kroppsfunktion. En

aktivitetsbegränsning kan i sin tur definieras som en begränsning eller svårighet som en individ innehar.

2.2 Webbtillgänglighet

Webbtillgänglighet innebär att alla skall ha tillgång till webben, oavsett nedsättning eller diverse fysiska eller psykiska hinder. Webbtillgänglighet gör det möjligt för personer med diverse nedsättningar att uppfatta och använda webben som den är. En tillgänglig webb gör det också möjligt för personer att utforska innehållet med hjälp av skärmläsare och annan hjälpteknik. Flera forskare poängterar att webbtillgänglighet, oavsett nedsättning, är en ytterst viktig aspekt**.** Webben bör med andra ord vara tillgänglig för alla och användbar av alla, oavsett av aktivitets- eller funktionsnedsättning. Personer med nedsatt syn upplever dock i dagens läge att webben är tre gånger så svår att använda jämfört med icke synskadade personer. Synskadade personer har därav en ytterst hög chans att uteslutas från webben.

World Wide Web Consortium skapade det första webbtillgänglighetsinitiativet i slutet av 1990-talet och dess mål var att förbättra tillgängligheten på webben för personer med olika funktionsnedsättningar. Initiativets vision var att skapa en internationell standard för att kunna definiera tillgängligheten av webbinnehåll. De första riktlinjerna var däremot inte allomfattande i och med att webben inte användes i lika stor grad som idag. Riktlinjerna har dock utökats i takt med digitaliseringen och idag innehåller de nya riktlinjerna regler och regleringar även för applikationer och diverse mobilt innehåll. Följaktligen är webbtillgänglighetsdirektiven även en ISO standard, som har godkänts av Europeiska unionens råd år 2016.

2.3 Användbarhet

Användbarhet har varit ett högt utnyttjat utvärderingsverktyg under en längre tidsperiod och har varit i bruk ända sen den första växelverkan mellan användare

och system. Användbarhet kan kännetecknas som en väsentlig del av hela designprocessen. Användbarhet definierar hur bra ett system kan användas av en specifik individ. Man kan vanligtvis påstå att en dålig användbarhet resulterar i att systemet förlorar sina potentiella kunder, medan en bra användbarhet lockar till sig nya kunder som gärna använder sig av systemet. Då det gäller användbarhet bland personer med diverse nedsättningar, påpekar tidigare studier att en betydande mängd pengar och fokus har investerats i webbutvecklare, medan fokusen på funktionshindrade användare lämnats bort.

När det gäller att mäta användbarhet bland personer med diverse nedsättningar, är det väsentligt att förstå personernas tankesätt samt avskilja de underliggande nedsättningarna som kan förekomma hos användarna. Enligt tidigare forskning, finns det i allmänhet fyra metoder för att utvärdera användbarheten av ett användargränssnitt; automatiska, empiriska, formella och informella metoder. Trots de olika metoderna för att mäta användbarhet, bör den mänskliga faktorn alltid vara närvarande.

3. Uppdragsgivare och PSOP-systemet

Kuntien Tiera Oy är ett företag, som ägs av över 300 kommunala organisationer. Företaget erbjuder produktiva IKT-tjänster för att stöda utvecklingen av den dagliga verksamheten hos sina kunder. Trots det breda utbudet av IKT-tjänster som tillhandahålls av företaget, fokuserar denna studie på systemet *Palveluseteli- ja ostopalvelujärjestelmä* (PSOP). PSOP-systemet är ett rikstäckande informationssystem, som gör det möjligt för kunden att jämföra priser, kvalitet och välja den mest lämpliga serviceproducenten åt sig själv. PSOP-systemet är Finlands populäraste system för att hantera servicesedlar och köpta tjänster.

4. Metod

Metoderna, som använts inom denna studie, följer metoderna för användbarhet. Den främsta orsaken för valet av dessa metoder, var att en mer teknisk undersökning av systemets tillgänglighetsfrågor har tidigare genomförts. Metoderna är en blandning av både kvalitativa och kvantitativa metoder. Målet med studien var att få en bättre förståelse och vidare utveckla ett bredare perspektiv på problemen relaterade till användbarhet- och tillgänglighet.
Som ett resultat av detta, utfördes ett användartest där deltagarna utförde tre uppgifter som var skapade av skribenten i förväg. Användartestet utfördes på både dator och på mobila enheter. Målgruppen för användartestet bestod av åtta blinda personer, som använde sig av diverse hjälpteknik. Även om användartesten utfördes på olika platser, både i Åbo och i Helsingfors, genomfördes arrangemangen med samma mål för att säkerställa att omgivningen inte skulle påverka resultaten. Utöver detta kan rekryteringsprocessen betraktas som den mest utmanande delen av studien, vilket resulterade i flera avgränsningar.
Följaktligen utfördes alla test under april månad. Varje test avspeglade en verklig växelverkan med systemet och som ett resultat av detta gav användartesten en enorm mängd data. Slutligen analyserades antalet upptäckta problem och deras svårighetsgrad väldigt noggrant.

Två ytterligare metoder användes för insamling av data. Online frågeformuläret *Google Forms*, användes för att skapa två frågeformulär. Det första frågeformuläret samlade in data gällande deltagarnas IT-kunskaper. Det andra frågeformuläret bestod av tio på förhand definierade frågor, vilket möjliggjorde räknandet av ett värde för systemets användbarhet (på engelska *System Usability Scale, SUS*). Man kan trots detta konstatera att användartestet var den metod, som bidrog med den mest värdefulla informationen.

5. Presentation av resultaten och analys av studien

Studien visar att PSOP-systemet har både användbarhets- och tillgänglighetsrelaterade problem och kan därför i viss utsträckning vara otillgänglig för personer med funktionsnedsättning. Från resultaten framgår det tydligt att även om ett system är på ett tekniskt sätt tillgängligt, kan ett användartest upptäcka hur otillgängligt systemet är och avslöja de verkliga problemen bland sina användare. I stora drag uppstod liknande problem på båda enheterna. De allvarliga problemen, som uppstod på båda enheterna, var bland annat inkonsekvent design, tillgång till bilder och tabeller och inmatningen av servicekupongen samt annan väsentlig information som hjälptekniken inte upptäckte.

Det andra frågeformuläret möjliggjorde räknandet av ett värde för systemets användbarhet (SUS). Kalkyleringen visade att PSOP-systemet har en användbarhet av 38,8 poäng på en skala från 0 till 100. I och med detta kan man konstatera att den uppfattade användbarheten av systemet är avsevärt låg och betydligt under genomsnittet för motsvarande system. Dock kan man även konstatera att deltagarna hade en varierande uppfattning av systemets användbarhet. Detta kom tydligt fram redan i användartesten, men säkerställdes av det andra frågeformuläret.

Sammanfattningsvis kan man konstatera att ingen av deltagarna kunde utföra alla uppgifter utan att stöta på några problem. Med hjälp av diverse metoder uppmanades deltagarna att ytterligare anstränga sig för att övervinna vissa problem. De ödesdigra problemen identifierades utan tvekan av systemet i uppgift 1 och 2. Däremot identifierades den minst felaktiga delen av systemet i uppgift 3.

6. Besvarandet av forskningsfrågorna

1: *Fastän ett tekniskt test gällande WCAG tidigare har genomförts, vilka problem stöter synskadade användare på när de använder sig av PSOP-systemet?*

Trots att en teknisk studie tidigare har genomförts, identifierades både tillgänglighets- och användbarhetsrelaterade problem genom att utföra ett användartest med åtta blinda deltagare. Ett stort antal olika problem framkom under testet, däremot varierade svårighetsgraden av problemen beroende på deltagarna och deras IT-kunskaper. Majoriteten av deltagarna identifierade dock liknande problem inom varje uppgift. Alla deltagare ansåg att systemet var inkonsekvent och enligt resultaten saknade systemet en ordentlig struktur. Utöver detta saknar systemet ett självförklarande innehåll, vilket återigen avspeglar både dålig användbarhethet och tillgänglighet.

2: *Vilka är orsakerna till de identifierade problemen inom PSOP-systemet och vilka åtgärder bör vidtas för att de skall elimineras?*

De identifierade problemen har relativt tydliga anledningar till att varför de inträffat. Dock är det emellertid relevant att inse att det kan trots allt finnas underliggande faktorer, som orsakat problemen. Flera av problemen är beroende på en otillräcklig design- och programmeringsprocess. Denna otillräckliga process resulterar i att systemet upplevs som inkonsekvent och där växelverkan drabbas avsevärt. En annan orsak bakom största delen av problemen är att systemet är högst antagligen planerat att användas av allmänheten. Problemen uppstår då innehållet inte är utformat enligt den logik eller det tankesätt som målgruppen innehar. I och med detta har planeringsprocessen förmodligen förbisett användningen av hjälpteknik, vilket resulterar i att vissa deltagare anser att systemet är i högsta grad otillgänglig.

Sammanfattningsvis kan man konstatera att även om PSOP-systemet anses vara tillgängligt på ett tekniskt sätt, kan ett användartest identifiera de problem som de verkliga användarna möter. Enligt analysen kan både användbarheten och tillgängligheten ökas genom att omprogrammera innehållet i enlighet med tankesättet av hela kundbasen, inklusive personer som lider av diverse nedsättningar.

7. Rekommendationer för Kuntien Tiera Oy

Det är ytterst viktigt för Kuntien Tiera Oy att ta hänsyn till problemen som studien lyft fram, eftersom PSOP-systemet vidareutvecklas efter årets andra kvartal. Denna studie har bevisat systemets verkliga användbarhet och tillgänglighetsgrad bland personer som använder sig av diverse hjälpteknik.
Efter en grundlig genomgång av problemen och ytterligare omprogrammering av systemet, bör Kuntien Tiera Oy utföra ett A/B-test och jämföra den gamla och den nya versionen med varandra för att identifiera ifall problematiken fortfarande kan upptäckas. Dessutom rekommenderas det att ett automatiserat test återigen utförs, för att säkerställa att tillgänglighetsdirektiven har uppfyllts. Det är anmärkningsvärt att det automatiserade testet bör alltid följas av ett användartest utfört av riktiga användare.

8. Förslag för vidare forskning

Det finns flera förslag för framtida forskning. Eftersom studien omfattar endast åtta deltagare, är det nödvändigt att utöka forskningsgruppen för att nå ett bredare resultat. En studie med en större forskningsgrupp resulterar troligtvis i intressanta resultat. Ett annat väsentligt forskningsförslag är att jämföra ett system mellan nybörjare och erfarna användare, för att kunna analysera ifall användarna identifierar liknande problem.
Följaktligen är det nödvändigt att poängtera att det andra frågeformuläret, vilket användes för att betygsätta användbarheten av systemet, inte nödvändigtvis är lämpligt för synskadade eller för personer med funktionsnedsättning. I och med detta bör framtida forskning identifiera ifall man på ett säkert sätt kan implementera frågeformuläret för denna målgrupp.

Slutligen kan man påstå att tillgänglighetsmetoder, framför allt automatiserade metoder, inte nödvändigtvis garanterar att ett system är tillgängligt i verkligheten. I och med detta bör vidare forskning utföras inom ämnet för att identifiera universella tillgänglighetsmetoder som säkerställer tillgänglighet för alla.

References

Atkins, N., Bennett, L., Domit, B., & Jones, J. (2011). *Heuristic evaluation*. Retrieved from http://booksite.mkp.com/barnum/testingessentials/pdfs/hufflepuff_he_report.pdf Retrieved: 05.03.2019

Bader, D., & Pagano, D. (2013). *Towards automated detection of mobile usability issues*. Retrieved from https://subs.emis.de/LNI/Proceedings/Proceedings215/341.pdf Retrieved: 19.05.2019

Bąkała, A., & Bąkała, M. (2014). *Web usability for people with disabilities. Studies & Proceedings of Polish Association for Knowledge Management,* Retrieved from https://pdfs.semanticscholar.org/83b5/eb64efb869d07e608550f775548a4cb93fb6.pdf Retrieved: 01.05.2019

Barnum, C. M. (2011). *Usability testing essentials*. Amsterdam: Elsevier. Retrieved from: http://digilib.stmik-banjarbaru.ac.id/data.bc/100.%20Other/2011%20Usability%20Testing%20Essentials.pdf Retrieved: 02.02.2019

Benyon, D. (2014). *Designing interactive systems* (3. ed. ed.). Harlow [u.a.]: Pearson. Retrieved from https://notendur.hi.is/mae46/Háskóli%20Íslands/2.%20misseri%20-%20Vor%202017/Viðmótsforritun/Bók/Designing.Interactive.Systems.A.Comprehensive.Guide.to.HCI.UX.and.Interaction.Design.3rd.edition_ibook4all.com.pdf Retrieved: 08.01.2019

Bevan, N., Carter, J., & Harker, S. (2015). *ISO 9241-11 revised: What have we learnt about usability since 1998?* Retrieved from https://pdfs.semanticscholar.org/6172/3cdd6f8195c8bb7407a04bb29a690de6a08c.pdf
Retrieved: 14.01.2019

Bradbard, D., & Peters, C. (2010). *Web accessibility theory and practice: An introduction for university faculty*. *The Journal of Educators Online, 7*(1) doi:10.9743/JEO.2010.1.1
Retrieved: 20.05.2019

Brajnik, G. (2008). *Beyond conformance: The role of accessibility evaluation methods*. *Web information systems engineering – WISE 2008 workshops* (pp. 63-80). Berlin, Heidelberg: Springer Berlin Heidelberg. doi:10.1007/978-3-540-85200-1_9
Retrieved: 20.02.2019

Brooke, J. (1996). SUS - A quick and dirty usability scale. Retrieved from https://hell.meiert.org/core/pdf/sus.pdf
Retrieved: 15.05.2019

Brooke, J. (2013). SUS: A retrospective. Retrieved from https://uxpa.org/sites/default/files/JUS_Brooke_February_2013_0.pdf
Retrieved: 06.05.2019

Brophy, P., & Craven, J. (2007). *Web accessibility*. Retrieved from https://www.ideals.illinois.edu/bitstream/handle/2142/3752/BrophyCraven554.pdf?sequence=2&sa=U&ei=noRiU7LaK8Oc2AWp_wE&ved=0CEQQFjAH&usg=AFQjCNE8-rgIDn5G5hH9KBPfsaakmg0oAQ
Retrieved: 20.01.2019

Budiu, R. (2017). *Quantitative vs. qualitative usability testing*. Retrieved from https://www.nngroup.com/articles/quant-vs-qual/
Retrieved: 05.02.2019

Cheng, L. C., & Mustafa, M. (2014). *A reference to usability inspection methods*. Paper presented at the doi:10.13140/2.1.2448.0964 Retrieved from https://search.datacite.org/works/10.13140/2.1.2448.0964 Retrieved: 20.02.2019

Council of the European Union. (2016). *Communication from the commission to the european parliament*. Retrieved from https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CONSIL:ST_11407_2016_INIT&from=DE Retrieved: 13.01.2019

Dix, A., Finlay, J., Abowd, G. D., & Beale, R. (2014). *Human–computer interaction* (Third edition ed.). New York: Pearson Education Limited. Retrieved: 15.01.2019

Eurostat. (2015a). Archive:*Disability statistics - barriers to social integration*. Retrieved from https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Archive:Disability_statistics_-_barriers_to_social_integration Retrieved: 20.01.2019

Eurostat. (2015b). Disability statistics background - *european health and social integration survey*. Retrieved from https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Disability_statistics Retrieved: 20.01.2019

Gui, H., Xu, Y., Bhasin, A., & Han, J. (May 18, 2015). *Network A/B testing*. Paper presented at the 399-409. doi:10.1145/2736277.2741081 Retrieved from http://dl.acm.org/citation.cfm?id=2741081 Retrieved: 19.05.2019

Gustafsson, H. (2016). *YK:N vammaissopimus käyttöön!* -käsikirja. Retrieved from https://www.invalidiliitto.fi/sites/default/files/2017-03/YK_Toolkit_2016_low.pdf Retrieved: 07.01.2019

Huang, C. J. (2003). *Usability of e-government web-sites for people with disabilities*. Paper presented at the 11 pp. doi:10.1109/HICSS.2003.1174330 Retrieved from https://ieeexplore.ieee.org/document/1174330
Retrieved: 21.05.2019

Hunter, M. (2010). *Usability testing by people with disabilities: Some guerrilla tactics*. Retrieved from http://uxpamagazine.org/guerilla_tactics/
Retrieved 03.03.2019

ISO 9241-11. (2018). *Ergonomic requirements for office work with visual display terminals (VDTs)* — Part 11: Guidance on usability. Retrieved from https://www.iso.org/obp/ui/#iso:std:iso:9241:-11:ed-2:v1:en
Retrieved: 21.02.2019

Mahlamäki, P. (2015). *YK:N yleissopimus vammaisten henkilöiden oikeuksista ja sopimuksen valinnainen pöytäkirja.* Suomen YK-liitto. Retrieved from https://www.ykliitto.fi/sites/ykliitto.fi/files/vammaisten_oikeudet_2016_net.pdf
Retrieved: 18.01.2019

Marcus, A. (2002). *Return on investment for usable user- interface design: Examples and statistics*. Retrieved from https://www.ida.liu.se/~TDDD26/material/economics5.pdf
Retrieved: 14.02.2019

Mariger, H. (2019). *Cognitive disabilities and the web: Where accessibility and usability meet?* Retrieved from http://ncdae.org/resources/articles/cognitive/
Retrieved: 10.04.2019

Moran, K. (2018). *Writing tasks for quantitative and qualitative usability studies.* Retrieved from https://www.nngroup.com/articles/test-tasks-quant-qualitative/
Retrieved: 01.03.2019

Morris, M. G., & Dillon, A. P. (1996). *The importance of usability in the establishment of organizational software standards for end user computing.* International Journal of Human - Computer Studies*, 45*(2), 243-258. doi:10.1006/ijhc.1996.0050
Retrieved: 17.05.2019

Nielsen Norman Group. (2010). *233 tips and tricks for recruiting users as participants in usability studies* (1st ed.) New Riders. Retrieved from http://proquestcombo.safaribooksonline.com/9780132107532
Retrieved: 21.03.2019

Nielsen Norman Group. (2014). *Turn user goals into task scenarios for usability testing*. Retrieved from https://www.nngroup.com/articles/task-scenarios-usability-testing/
Retrieved: 03.03.2019

Nielsen, J. (1994). *Usability inspection methods*. John Wiley and Sons, New York, 25-62.
Retrieved: 01.02.2019

Nielsen, J. (2000). *Why you only need to test with 5 users*. Retrieved from http://www.mathcs.emory.edu/~cengiz/cs540-485-soft-eng-fa14/resources/5userTesting.pdf
Retrieved: 26.05.2019

Nielsen, J. (2010). *Defining usability*. (2010 Elsevier, Inc) Retrieved from https://booksite.elsevier.com/samplechapters/9780123751140/02~Chapter_1.pdf
Retrieved: 13.01.2019

Nielsen, J. (2012). *Usability 101: Introduction to usability.* Retrieved from https://www.nngroup.com/articles/usability-101-introduction-to-usability/
Retrieved: 26.05.2019

Nielsen, J., & Molich, R. (1990). *Heuristic evaluation of user interfaces*. Retrieved from http://cs.ashoka.edu.in/cs102/papers/heuristic-evaluation-of-user-interfaces-nielsen.pdf
Retrieved: 14.02.2019

Norman, K. L., & Panizzi, E. (2006). *Levels of automation and user participation in usability testing*. Interacting with Computers*, 18*(2), 246-264. doi:10.1016/j.intcom.2005.06.002
Retrieved: 22.02.2019

Orfanou, K., Tselios, N., & Katsanos, C. (2015). *Perceived usability evaluation of learning management systems: Empirical evaluation of the system usability scale*. *The International Review of Research in Open and Distributed Learning, 16*(2) doi:10.19173/irrodl.v16i2.1955
Retrieved: 17.05.2019

Petrie, H., & Bevan, N. (2009). *The evaluation of accessibility, usability and user experience.* Retrieved from https://pdfs.semanticscholar.org/6896/17fb8d66ec0d3cd2e1012171461c22f4c6f5.pdf
Retrieved: 13.02.2019

Petrie, H., & Kheir, O. (Apr 29, 2007). *The relationship between accessibility and usability of websites.* Paper presented at the 397-406. doi:10.1145/1240624.1240688 Retrieved from http://dl.acm.org/citation.cfm?id=1240688
Retrieved: 20.02.2019

Power, C., Freire, A., Petrie, H., & Swallow, D. (May 5, 2012). *Guidelines are only half of the story*. Paper presented at the 433-442. doi:10.1145/2207676.2207736 Retrieved from http://dl.acm.org/citation.cfm?id=2207736
Retrieved: 07.01.2019

Reid, L., & Snow-Weaver, A. (Apr 21, 2008). *WCAG 2.0*. Paper presented at the 109-115. doi:10.1145/1368044.1368069 Retrieved from http://dl.acm.org/citation.cfm?id=1368069
Retrieved: 07.01.2019

Richards, J., & Hanson, V. (May 17, 2004). *Web accessibility*. Paper presented at the 72-79. doi:10.1145/988672.988683 Retrieved from http://dl.acm.org/citation.cfm?id=988683
Retrieved: 08.01.2019

Rukshan, A., & Baravalle, A. (2012). *Automated usability testing: Analysing asia web sites*. Retrieved from https://arxiv.org/abs/1212.1849
Retrieved: 01.02.2019

Schneiderman, B. (2000). *Universal usability*. Retrieved from http://www.cs.umd.edu/~ben/p84-shneiderman-May2000CACMf.pdf
Retrieved: 03.02.2019

Sevilla, J., Herrera, G., Martínez, B., & Alcantud, F. (2007). *Web accessibility for individuals with cognitive deficits: A comparative study between an existing commercial web and its cognitively accessible equivalent*
doi:10.1145/1279700.1279702
Retrieved: 09.01.2019

Sutcliffe, A., & Gault, B. (2004). *Heuristic evaluation of virtual reality applications*. *Interacting with Computers, 16*(4), 831-849. doi:10.1016/j.intcom.2004.05.001
Retrieved: 14.03.2019

Termens, M., Ribera, M., Porras, M., Boldú, M., Sulé, A., & Paris, P. (Apr 20, 2009). *Web content accessibility guidelines*. Paper presented at the 1171-1172. doi:10.1145/1526709.1526912 Retrieved from http://dl.acm.org/citation.cfm?id=1526912
Retrieved: 17.01.2019

United Nations. (2006). *Convention on the rights of persons with disabilities and optional protocol*. Retrieved from http://www.un.org/disabilities/documents/convention/convoptprot-e.pdf
Retrieved: 04.01.2019

Valtiovarainministeriö. (2018). *Saavutettavuus*. Retrieved from https://vm.fi/saavutettavuusdirektiivi
Retrieved: 23.01.2019

van der Geest, T. (2006). Conducting usability studies with users who are elderly or have disabilities. *Technical Communication, 53*(1), 23-31. Retrieved from http://www.ingentaconnect.com/content/stc/tc/2006/00000053/00000001/art00004
Retrieved: 13.02.2019

W3C. (1999). *Web content accessibility guidelines 1.0*. Retrieved from https://www.w3.org/TR/WCAG10/#Guidelines
Retrieved: 15.01.2019

W3C. (2005). *Web content accessibility guidelines (WCAG) overview.* Retrieved from https://www.w3.org/WAI/standards-guidelines/wcag/#iso
Retrieved: 15.01.2019

W3C. (2008). *Web content accessibility guidelines (WCAG) 2.0*. Retrieved from https://www.w3.org/TR/WCAG20/
Retrieved: 17.01.2019

W3C. (2014). *Accessibility testing*. Retrieved from https://www.w3.org/wiki/Accessibility_testing#User_testing
Retrieved: 01.02.2019

W3C. (2018). *Web content accessibility guidelines (WCAG) 2.1.* Retrieved from https://www.w3.org/TR/WCAG21/
Retrieved: 07.02.2019

WAI. (2005). *Introduction to web accessibility*. Retrieved from https://www.w3.org/WAI/fundamentals/accessibility-intro/ Retrieved: 09.01.2019

Whiteside, J. B., & Holtzblatt, K. (1988). *Usability engineering: Our experince and evalution*. *Handbook of human-computer interaction – Chapter 36* (pp. 791-817) Elsevier B.V. doi:10.1016/B978-0-44-70536-5.50041-5 Retrieved: 15.01.2019

Wilson, C. (2014). *User interface inspection methods.* US: Morgan Kaufmann Publisher Inc. Retrieved from https://ebookcentral.proquest.com/lib/[SITE_ID]/detail.action?docID=1568337 Retrieved: 18.03.2019

World Health Organization, & World Bank. (2011). *Summary: World report on disability 2011* Geneva: World Health Organization. Retrieved from https://apps.who.int/iris/bitstream/handle/10665/70670/WHO_NMH_VIP_11.01_eng.pdf?sequence=1&isAllowed=y Retrieved: 25.01.2019

Zhang, Z., Basili, V., & Schneiderman, B. (1999). *Perspective-based usability inspection: An empirical validation of efficacy. Empirical Software Engineering,* 4(1), 43-69. *doi:1009803214692.* Retrieved from https://link.springer.com/article/10.1023/A:1009803214692 Retrieved: 23.02.2019

Appendix A: *Advertisement*

Hyvä vastaanottaja,

Nimeni on Oscar Ranta ja olen tietojärjestelmätieteen opiskelija Åbo Akademissa ja teen maisterintutkielmaani Palveluseteli- ja ostopalvelujärjestelmän käytettävyydestä. Toimeksiantajana toimii kuntaorganisaatioiden omistama inhouse-yhtiö, Kuntien Tiera Oy. Toimeksiannon ohjaajana toimii Kuntien Tiera Oy:n Jani Grönberg. Valtakunnallisena, kuntaorganisaatioiden omistamana yhtiönä Kuntien Tiera Oy tarjoaa laadukkaita ja tuotteistettuja ICT-palveluita asiakkailleen, sujuvan arjen ja toiminnan kehittämisen tueksi.

Palveluseteli- ja ostopalvelujärjestelmä (PSOP) on Suomen suosituin järjestelmä, jonka verkkopalvelussa asiakas voi tarkastella myönnettyjä palveluseteleitä ja käytettävissä olevia palvelun tuottajia. PSOP-järjestelmän avulla asiakas pystyy vertailemaan hintoja ja laatua sekä valitsemaan itselleen sopivimman palveluntuottajan. PSOP-järjestelmä mahdollistaa valinnanvapautta eri palveluiden valitsemisessa ja soveltuu käytettäväksi kunnasta ja palvelusta riippumatta.

Etsin suurennus- ja/tai ruudunlukuohjelmia tai muuta avustavaa teknologiaa käyttäviä testihenkilöitä, jotka ovat saaneet PSOP-järjestelmän kautta palvelusetelin. Testihenkilöiltä ei vaadita erityisiä valmiuksia. Sekä aloittelijat että edistyneet käyttäjät voivat osallistua testaukseen. Aikaisempi kokemus PSOP-järjestelmästä on suotavaa mutta myös vähempi kokemus riittää. Käytettävyystesti kestää noin 30 minuuttia ja tutkielmassa ei kerätä asiakastietoja. Tutkielma on anonyymi.

Testaus tehdään Internet sivulla tietokoneen selaimella. Testihenkilöt voivat itse valita missä testi suoritetaan. Se voidaan tehdä esimerkiksi Kuntien Tiera Oy:n tiloissa Turussa, osallistuvan henkilön luona tai osallistuvan henkilön itse valitsemassa paikassa. Testaus voidaan siis suorittaa paikkakunnasta riippumatta testihenkilön toiveiden mukaisesti.

Testihenkilö suorittaa ennalta määriteltyjä tehtäviä, jotka kuvastavat PSOP-järjestelmän jokapäiväistä toimintaa. Tehtävät ovat avoimia ja tehtäviin ei ole oikeita vastauksia. Pyrin kartoittamaan ongelmat, johon käyttäjät törmäävät.

Käytettävyystestauksen pohjalta pyritään vastaamaan tutkimuskysymykseen ja arvioimaan kuinka hyvin PSOP-järjestelmä on huomioinut saavutettavuusdirektiivin kriteerit ja miten se näkyy käytettävyydessä. Tutkimustuloksen pohjalta pyritään lisäämään PSOP-järjestelmän asiakastyytyväisyyttä sekä huomioimaan ja mahdollisesti parantamaan havaittuja ongelmia. Tutkimus toimii pohjana PSOP-järjestelmän jatkokehitykselle.

Mikäli kiinnostuksenne heräsi ja olette halukkaita osallistumaan käytettävyystestiin, voi ilmoittautumisen lähettää sähköpostitse osoitteeseen

oscar.ranta@gmail.com

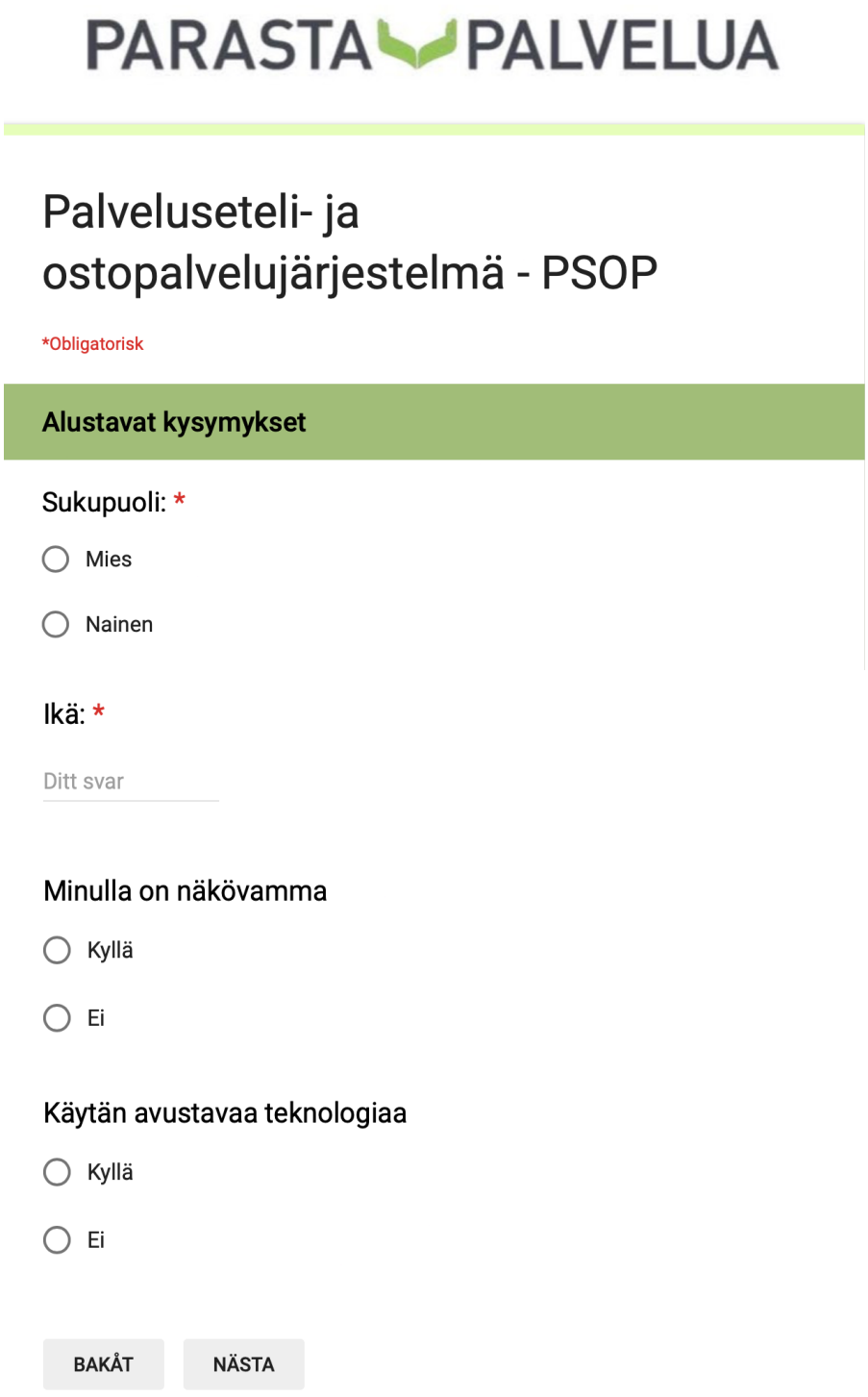
Kiitos

Oscar Ranta
Åbo Akademi

Kuntien Tiera Oy | **www** www.tiera.fi | **Y-tunnus** 2362180-3
Tammasaarenkatu 3 | **puhelin** 020 734 7777 | **ALV-numero** FI23621803
00180 Helsinki | **sähköposti** info@tiera.fi | **Kotipaikka** Helsinki

1/1

Appendix B: *Pre-test questionnaire*



PARASTA PALVELUA

Palveluseteli- ja ostopalvelujärjestelmä - PSOP

*Obligatorisk

Alustavat kysymykset

Sukupuoli: *

- Mies
- Nainen

Ikä: *

Ditt svar

Minulla on näkövamma

- Kyllä
- Ei

Käytän avustavaa teknologiaa

- Kyllä
- Ei

BAKÅT NÄSTA

Palveluseteli- ja ostopalvelujärjestelmä - PSOP

*Obligatorisk

Tietotekninen osaaminen

Arvioi tietotekninen osaamisesi: *

	1	2	3	4	5	
Aloittelija	○	○	○	○	○	Edistynyt käyttäjä

Kuinka usein käytät tietokonetta tai muuta vastaavaa laitetta?

- ○ Päivittäin
- ○ Viikoittain
- ○ Kuukausittain
- ○ En juuri koskaan tai hyvin harvoin

Tunnit päivässä, joka on käytetty verkossa

- ○ Alle tunti
- ○ 1-2 tuntia
- ○ 2-4 tuntia
- ○ 4-6 tuntia
- ○ Enemmän kuin 6 tuntia

Mitä laitetta käytät mieluiten?

- ○ Tietokonetta
- ○ Mobiililaitetta

Kohtaatko yleensä ongelmia tai esteitä internettiä selatessa?

- ○ En
- ○ Övrigt: _____

Appendix C: *Test tasks*

PSOP JÄRJESTELMÄN KÄYTETTÄVYYSTESTAUS

1. **HAMMAS SKENAARIO**

*Sinua on jo pidemmän aikaa **häirinnyt hammaskipu** ja olet nyt päättänyt, että asialle pitää tehdä jotakin. Olet tullut siihen lopputulokseen, että **hammas pitää poistaa**.*

*Olet äskettäin kuullut, että verkkopalvelu nimeltään "Palveluseteli- ja ostopalvelujärjestelmä" mahdollistaa **suun terveydenhuolto** palveluiden kilpailutuksen ja tämän takia päätät käyttää palvelua löytääksesi oikean toimenpiteen. **Turun kaupunki** on myöntänyt sinulle myös palvelusetelin, jonka tiedot päätät syöttää "Palveluseteli- ja ostopalvelujärjestelmään".* Palveluseteleitä on myönnetty yksi kappale, arvoltaan 65 €.

*Haluat nyt löytää itsellesi sopivimman palveluntuottajan, joka tarjoaa haluamaasi palvelua. Olet vakuuttunut siitä, että kyseessä on ainoastaan yksi oikea toimenpide (eli **hampaan poisto)**. Olet kuullut, että muut palvelunsaajat ovat arvostelleet palveluita "Palveluseteli- ja ostopalvelujärjestelmän" avulla ja tämän takia aiot valita palvelun, joka on saanut parhaat arvostelut.*

Seuraavaksi luettelen sinulle tiedot, joita tarvitset oikean palvelun löytämiseksi:

- Sinulle on myönnetty 1kpl palveluseteleitä, arvoltaan 65€
- Et missään nimessä halua jonottaa saadaksesi palvelua
- Valitse **kieli**, jolla haluat palvelua ja tunnista itsesi **kohderyhmästä** sekä määritä **esteettömyys** tarpeittesi mukaisesti
- Valitse **vertailuun kaksi tai useampi palveluntuottaja**

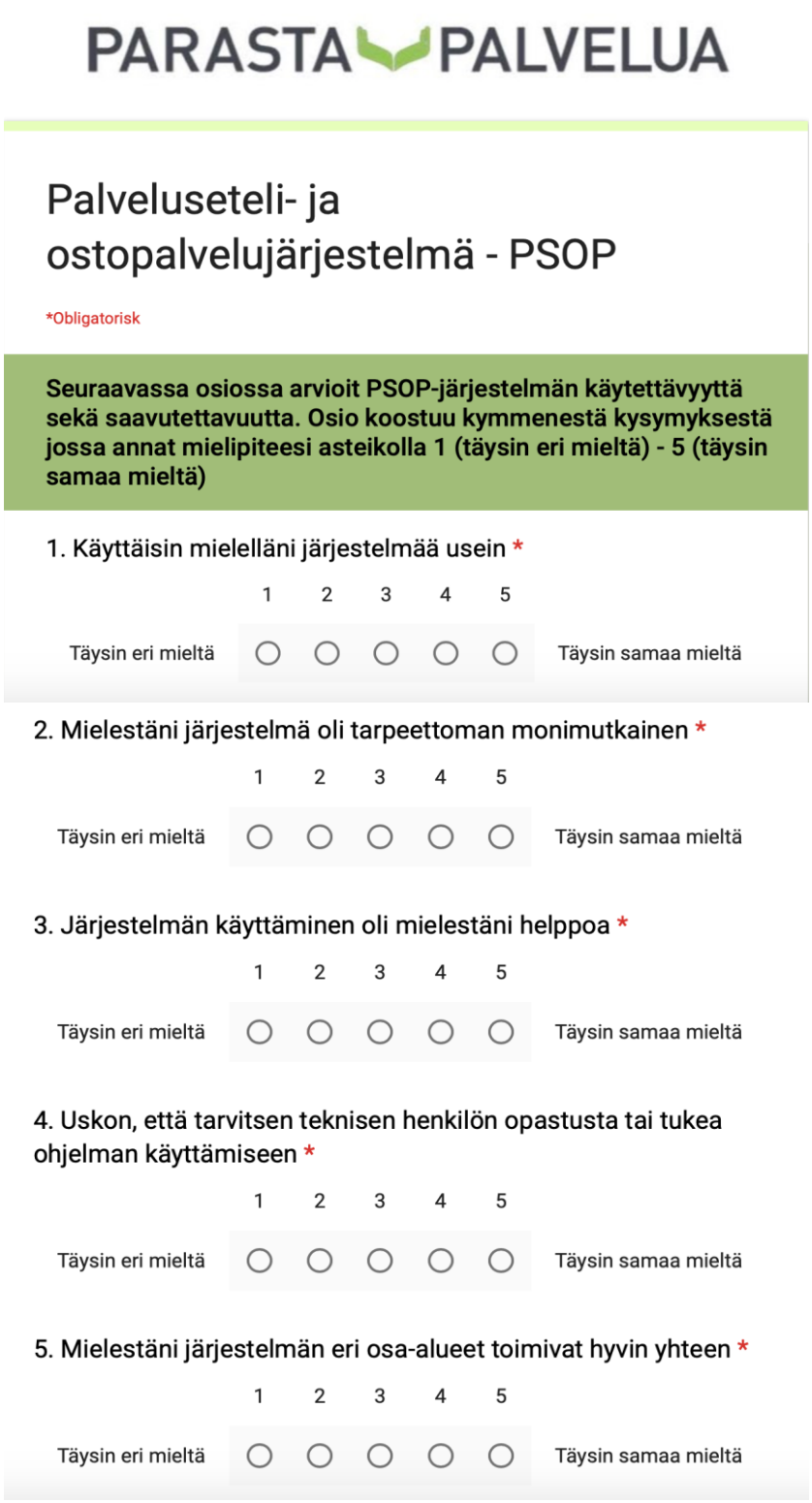
2. Palvelun arviointi

Etsi sinulle myönnetty palvelu ja täytä palvelun arviointi lomake haluamasi mukaisesti.

3. Palautteen anto

Haluat antaa palautetta saamallesi palvelulle. Saat itse valita mille yritykselle ja palvelulle haluat antaa palautetta. Palautteen muun sisällön saat myös vapaasti valita.
Ainoa kriteeri on, että haluat varmistaa, että yritys tai työntekijä (palveluntuottaja) antaa vastauksen sinun lähettämällesi palautteelle. Sinun tulee myös osoittaa että koko palaute on lähetetty onnistuneesti.

Appendix D: *Post-test questionnaire (SUS)*



PARASTA PALVELUA

Palveluseteli- ja ostopalvelujärjestelmä - PSOP

*Obligatorisk

Seuraavassa osiossa arvioit PSOP-järjestelmän käytettävyyttä sekä saavutettavuutta. Osio koostuu kymmenestä kysymyksestä jossa annat mielipiteesi asteikolla 1 (täysin eri mieltä) - 5 (täysin samaa mieltä)

1. Käyttäisin mielelläni järjestelmää usein *

	1	2	3	4	5	
Täysin eri mieltä	○	○	○	○	○	Täysin samaa mieltä

2. Mielestäni järjestelmä oli tarpeettoman monimutkainen *

	1	2	3	4	5	
Täysin eri mieltä	○	○	○	○	○	Täysin samaa mieltä

3. Järjestelmän käyttäminen oli mielestäni helppoa *

	1	2	3	4	5	
Täysin eri mieltä	○	○	○	○	○	Täysin samaa mieltä

4. Uskon, että tarvitsen teknisen henkilön opastusta tai tukea ohjelman käyttämiseen *

	1	2	3	4	5	
Täysin eri mieltä	○	○	○	○	○	Täysin samaa mieltä

5. Mielestäni järjestelmän eri osa-alueet toimivat hyvin yhteen *

	1	2	3	4	5	
Täysin eri mieltä	○	○	○	○	○	Täysin samaa mieltä

6. Mielestäni järjestelmässä oli liian paljon epäjohdonmukaisuutta *

	1	2	3	4	5	
Täysin eri mieltä	○	○	○	○	○	Täysin samaa mieltä

7. Uskon, että useimmat oppivat järjestelmän käytön erittäin nopeasti *

	1	2	3	4	5	
Täysin eri mieltä	○	○	○	○	○	Täysin samaa mieltä

8. Koin järjestelmän käytön erittäin hankalaksi *

	1	2	3	4	5	
Täysin eri mieltä	○	○	○	○	○	Täysin samaa mieltä

9. Tunsin itseni hyvin varmaksi, kun käytin järjestelmää *

	1	2	3	4	5	
Täysin eri mieltä	○	○	○	○	○	Täysin samaa mieltä

10. Minun piti opetella useita asioita ennen järjestelmän käytön aloittamista *

	1	2	3	4	5	
Täysin eri mieltä	○	○	○	○	○	Täysin samaa mieltä

Tässä osiossa voit vielä vapaasti antaa muuta palautetta tai huomioitavaa PSOP-järjestelmästä

Ditt svar