Understanding accessibility problems of blind users on the web

Andreas Savva

Doctor of Philosophy

University of York
Computer Science

September 2017

Abstract

The web is an eminently visual medium. However, not everyone accesses web content visually. Research shows that using the web is challenging for blind users. To create a good user experience for blind users on the web, we need a comprehensive understanding of the users' problems. Currently, there is little knowledge about the problem differences between blind and sighted users, which makes it difficult to suggest and test design solutions that address these problems.

This research aims to provide a further understanding of the problems blind users have on the web by comparing and contrasting problems between blind and sighted users and testing how design solutions to prevalent problems benefit blind users' experience.

The first study draws together the research literature into a common unified definition of web accessibility that was used to operationalise studies. The second study compared which verbal protocol (concurrent or retrospective) is better in user-based studies. The results showed that retrospective verbal protocol is a better option for eliciting problems on the web for blind and sighted users.

Then, an empirical study compared the problems between blind and sighted users on the web. The results showed that the problems the two user groups encounter largely differ. There are specific problem types distinct to blind users, but also the characteristics of the problem types that had instances by both user groups were very different. Moreover, many problems blind users encounter were in relation to the search and browse features of the websites.

A further investigation by two studies with blind users of how specific design solutions to prevalent problems users had (poor page structure, lack of feedback and excessive effort) in this specific design aspect showed that simple design solutions improve specific aspects of users' experience. Although, for major improvements in the overall user experience a combination of design solutions is needed.

Table of Contents

List of Tables

List of Figures

Acknowledgements

I would like to take the opportunity and thank all the people who help me across this long journey and contributed to the completion of this research.

First, I would like to express my sincere gratitude to my supervisors Professor Helen Petrie and Dr. Christopher Power for the continuous support of my thesis work, for their patience, motivation and immense knowledge. Their guidance helped me in all of my time during this research program.

Also, I would like to thank my assessors Dr. Paul Cairns and Dr. David Sloan for their valuable contribution to the evaluation of this thesis.

My thanks to all participants that took part in the study. I would also like to thank all my friends and colleagues at the University of York.

Thanks to the Engineering and Physical Science Research Council of UK and the Cyprus State Scholarship Foundation for the financial support to develop this research thesis.

My special thanks go to my family, to my parents and my sister, who were always there for me during my studies.

Declaration

I declare that the research work presented in this thesis, if not otherwise stated, is my own. This work has not been submitted for any other degree at this or any other university. All sources are acknowledged as references. All empirical studies were approved by the University ethics committee. Parts of this thesis have been published in the following conferences:

Petrie, H., **Savva, A.**, & Power, C. (2015). Towards a unified definition of web accessibility. In: *Proceedings of the 12th Web for all Conference*, 35.

Savva, A., Petrie, H., & Power, C. (2015). Comparing concurrent and retrospective verbal protocols for blind and sighted users. In: *Human-computer interaction*, 55-71.

Savva, A., Petrie, H., & Power, C. (2016). Types of Problems Elicited by Verbal Protocols for Blind and Sighted Participants. In: *International Conference on Computers Helping People with Special Needs*, 560-567.

Chapter 1. Introduction

The web offers instant access to an unprecedented amount of information and many important services. The fast evolution of the web made it an important medium for being an active participant in society. People can manage their bank accounts, pay their bills and taxes, communicate with friends, book flight tickets, buy groceries and clothes, and even complete university degrees. Many people can complete daily activities on the web, without having to leave their house or wait in long queues. The vast development of the web made it an indispensable part of peoples' everyday lives (Hoffman, Novak, & Venkatesh, 2004).

People with disabilities can have many benefits from being able to access and use information on the web. Without leaving their homes, people can independently have access to information and services that are available on the web. For example, blind users can access information the same time it becomes available to sighted users, without having to wait for the information to become available in alternative formats, for instance, Braille and audio reading materials. People with limited manual dexterity, that are not able to use the keyboard or the mouse of the computer, can use speech recognition software to access the information and services, which before did not seem feasible without assistance. Many services and activities that were difficult or impossible to be performed by people with disabilities, independently, can now be done on the web often using assistive technologies.

There is still a lot to understand about how to create a good accessible user experience for blind users on the web, which is what this thesis addresses.

1.1. Background to the research

Over a billion people, around 15% of the world's population have some form of disability, with the number growing due to population ageing (World Health Organization, 2016). Although accessing and using the web can have many benefits to people with disabilities, it appears to be a real challenge for them to use it (Disability Rights Commission, 2004; André Pimenta Freire, 2012).

People with disabilities have every right to access and use the websites. Making websites accessible to everyone is, first of all, a moral obligation. Some countries, to protect people with disabilities from discrimination, have legislation that requires websites to be accessible to everyone. For instance, the Section 508 of the Rehabilitation Act in US and the Equality Act (2010) in the UK. From a business point of view, having an accessible website means expanding the target market, which can lead to more consumers and enhance a company's reputation. The number of disabled people in the UK is approximately 10 million people, which accounts for 15% of the population, with their spending power estimated at 212 billion a year (Department for Work and Pensions Office for Disability Issues and The Rt Hon Mark Harper, 2014). Also, if websites are not accessible, then many people with disabilities will be dependent on others, which can result in a loss of confidentiality, as they will be dependent on others to read personal information (e.g. bank details, medical information).

Even though insights of the difficulties people with disabilities have on the web were raised through early research (Oppenheim & Selby, 1999), it seems that there has not been any improvement through the development of the web over the years (Disability Rights Commission, 2004; André Pimenta Freire, 2012). Websites that are not designed to accommodate the needs of the broadest range of users can create several challenges for people with disabilities (Coyne & Nielsen, 2001; Disability Rights Commission, 2004; André Pimenta Freire, 2012; Petrie & Kheir, 2007).

When considering the different disabled user groups, research demonstrated that blind users, whose number amounts to 36 million people (World Health Organization, 2017), face the most difficulties using the web (Disability Rights Commission, 2004; André Pimenta Freire, 2012). Extant research recognises that blind users encounter many problems on the web, yet no research provides an extensive detailed analysis of the characteristics of the problems they encounter over sighted users. As most websites are mainly designed on users' ability to cope with information visually, it is important to understand the differences of the problems on the web between blind and sighted users in order to design websites that take into account the different user needs and create similar experiences. Petrie and Kheir (2007) demonstrated that the problems blind and sighted users have largely differ with very few problems being encountered by both user groups. However, the study did not provide any further information about the problem differences and similarities between the two user groups.

While there are known differences as well as known overlaps in the problems blind and sighted users encounter on the web, there is little knowledge in the literature about these problems as well as the causes of these problems. The limited knowledge of the similarities and differences of the problems encountered on the web between blind and sighted users makes it difficult to design websites that accommodate the needs of blind users. As previous research showed that the existing accessibility techniques are not sufficient to cover all the problems blind users have on the web (André Pimenta Freire, 2012; Power, Freire, Petrie, & Swallow, 2012), designers and developers work became even more difficult for creating a good accessible user experience.

There is a need to design solutions that properly address the problems blind users have and create a better experience on the web. This can be done via a thorough understanding the problems that are distinct to blind users as well as the problems that are shared with sighted users. Then, design solutions can be proposed that properly address these problems and improve users' experience. However, to be able to suggest design solutions to the key problems blind users encounter on the web, a thorough understanding of how specific design solutions can benefit blind users' experience on the web is required.

This research project, therefore, provided an important opportunity to advance our understanding of the problems blind users have on the web by expanding our knowledge of the problem similarities and differences between blind and sighted users, as well as to enhance our understanding of how specific design solutions benefit users' experience.

Before conducting studies that investigate the problems blind users have on the web, there are two important considerations that should not be overlooked. First, there is no agreement on the definition of web accessibility in the community. This inability of a universally accepted definition makes it difficult to design studies that properly study the concept. A consensus on the definition of web accessibility is required in order to talk cohesively about the concept and be clear about what is measured in relation to the concept. The second consideration that should not be pass unseen is that almost all studies with blind users that look into the problems they encounter on the web are conducted with users thinking aloud as they were performing the tasks. Although thinking aloud as users do the tasks seems the norm, researchers raised their concerns for this approach (Chandrashekar, Stockman, Fels, & Benedyk, 2006; Coyne & Nielsen, 2001), as it might not be the most appropriate method for blind users. Alternative approaches should also be considered as they may be better for eliciting problems with blind users on the web. Thus, these two considerations must first be looked before studies are conducted that investigate the main aims of this research project.

1.2. Research questions and objectives

The overall goal of this programme of research in this thesis is to contribute into the areas of accessibility by investigating the problems blind users have on the web and design solutions for some of the most prevalent problems. The research questions that the research aims to answer are:

- *What are the problem similarities and differences between blind and sighted users on the web?*
- *What are the benefits of specific design solutions to the key problems on blind users' experience?*

The objectives were:

- To compare and contrast problems between blind and sighted users.
- To elicit and classify the range and diversity of user problems.
- To provide details of what causes the user problems.
- To relate how common key problems influence the effectiveness and efficiency of users and the perceived usability of a website.

- To evaluate how specific design solutions can address the key problems blind users have to improve the effectiveness and efficiency of users and the perceived usability of a website.
- To investigate whether simple design solutions to the problems are enough to improve blind users' experience on the web.

During the course of this investigation, it was determined that there are two major gaps in the existing research that hindered the experimental research in this area. First, there was lack of clarity on a definition of web accessibility, which makes it difficult to operationalise experiments and investigate user problems and their potential solutions. Second, there is no research that investigates which verbal protocol can be considered a better option for eliciting problems on the web for blind users. Therefore, the following research questions were proposed to address these issues before undertaking the core of the programme of research:

- *What are the most frequent components that researchers consider as part of the concept of web accessibility?*
- *Which verbal protocol can be considered a better option for user-based studies with blind and sighted users on the web?*

The objectives were:

- To undertake an analytical study that draws together the research literature into a common unified definition of web accessibility.
- To evaluate which verbal protocol can be considered a better option for eliciting problems on the web for blind and sighted users.

1.3. Outline of the structure of the thesis

The overall structure of the thesis takes the form of nine chapters, including this introductory chapter. Chapter Two presents a review of the literature. It starts by looking into different definitions of web accessibility and focusing on the studies that look into the problems blind users have on the web.

The third chapter begins by describing a qualitative analysis of different definitions of web accessibility. It includes the sampling method, analysis, and it draws together a unified definition of web accessibility that encompasses all the core complements of the concept that are considered by researchers. The proposed unified definition of web accessibility was used for the operationalisation of the experiments of the next studies of this thesis work.

Chapter Four presents an investigation of which of the two verbal protocols (concurrent verbal protocol and retrospective verbal protocol) is better for eliciting problems on the web for blind and sighted users. The results of this study guided the verbal protocol used in the next of this thesis.

Chapter Five presents an investigation of the problem similarities and differences between blind and sighted users on the web. The study revealed that the characteristics of the problem types the two user groups encounter largely differ, with a number of distinct problems for blind users. Also, many of the problems blind users encounter were in relation to the search and filtering browsing of content.

In Chapter Six, an analytical study of the design features between the websites used in the previous study and a similar type of websites is presented to ensure the generalisability of the results of the previous study. Based on the results, many of the problems found by blind users in the study in Chapter Five can be generalised to other shopping websites, as there were many commonalities between the design features of the websites used in the previous study and other shopping websites.

Chapter Seven proposes specific design solutions to some of the key problems blind users have when searching and browsing through the website content.

Chapter Eight presents an investigation of the specific design solutions to the problems blind users have to check the benefits on users’ experience.

Chapter Nine presents a confirmative study of whether the results found in the previous study of this thesis maintain when users do an exploratory task that has higher ecological validity.

The final chapter presents the conclusion of the studies by summarising the main contributions of the work in this thesis and list areas that worth further investigation in future work.

Chapter 2. Literature Review

2.1. What is Web Accessibility

Access by everyone is one of the primary motivations of the creation of the world wide web (Berners-Lee, 1997). Research investigating the accessibility of web content has been a topic of considerable importance since its early stages (Paciello, 1996a, 1996b, 1996c).

Even though web accessibility has been widely studied during the last years, different definitions of the concept can be found in the literature. There are definitions from standards issued by international bodies, such as the Web Accessibility Initiative (WAI, 2005) of the World Wide Web Consortium (W3C) and the International Standards Organization (ISO 9241-171, 2008), and government bodies such as the Section 508 of the Rehabilitation Act of the USA and the British Standards Organization (BSI, 2010). Some researchers relate accessibility to other concepts, such as usability (Petrie & Kheir, 2007; Shneiderman, 2002; Thatcher et al., 2002) and user experience (Aizpurua, Harper, & Vigo, 2016; Horton & Quesenbery, 2014).

2.1.1. Accessibility from the point of view of Standards

The accessibility definitions from standards are presented in Table 1. As can be seen, each definition has a different scope and nature. Some definitions refer to different levels of interaction, some refer to all people, some refer to usability properties such as effectiveness, efficiency and satisfaction. However, what all definitions from standards have in common is that they all refer to people with disabilities.

Table 1. Web accessibility definitions from Standards.

Source	**Definition**
Section 508 of the Rehabilitation Act	Technology is accessible if it can be used as effectively by people with disabilities as by those without.
WAI (2005)	Web accessibility means that people with disabilities can use the Web. More specifically, Web accessibility means that people with disabilities can perceive, understand, navigate and interact with the Web, and that they can contribute to the Web.
ISO 9241-171 (2008)	... the usability of a product, service, environment or facility by people with the widest range of capabilities.
BSI (2010)	… usability of a product, service, environment or facility by people within the widest range of capabilities. The concept of accessibility addresses the full range of user capabilities and is not limited to users who are formally recognized as having disability. The usability-oriented concept of accessibility aims to achieve levels of effectiveness, efficiency and satisfaction that are as high as possible considering the specified context of use, while paying attention to the full range of capabilities within the user population. In a web context, accessibility means the degree to which people with disabilities can

	perceive, understand, navigate, and interact with the web, and that they can contribute to the web.

The Web Accessibility Initiative (WAI) of the World Wide Web Consortium (W3C) propose a definition of web accessibility that refers to different levels of interaction with the web content, such as "perceive, understand, navigate and interact". However, the criteria promoted to evaluate the accessibility of a website are not user-based. WAI (2005) published a set of guidelines to help developers and designers create more accessible websites (W. Chisholm, Vanderheiden, & Jacobs, 1999). The first version of guidelines was published in 1999, the Web Content Accessibility Guidelines 1.0 (WCAG 1.0). However, as the web started evolving rapidly, W3C recognised that WCAG 1.0 would be outdated, thus in 2008 the second version of WCAG (WCAG 2.0) was published (Caldwell, Cooper, Guarino Reid, & Vanderheiden, 2008). W3C announced plans for the development of an updated version of the guidelines, the WCAG 2.1 by mid-2018. In addition, W3C is planning the development of a major update of the guidelines, version 3.0, which will incorporate many of the changes that will be introduced in WCAG 2.1 (Cooper, 2016). According to W3C, this set of guidelines can be used as criteria for evaluating and achieving web accessibility.

Section 508 of the Rehabilitation Act of 1973 requires all US Federal agencies and entities to make their electronic and information technology accessible to disabled people (Rutter et al., 2007). The definition of accessibility proposed by Section 508 refers to equal access for people with disabilities. Even though the definition refers to the effective use, which may imply the evaluation of websites with usability qualities, Section 508 requires websites to conform to technical standards. Initially, Section 508 had their own set of guidelines, with the majority of them being based on WCAG 1.0 guidelines. A recent update of Section 508 standard requires the conformance of websites to WCAG 2.0 guidelines.

The International Organisation for Standardisation (ISO 9241-171, 2008) provides guidance and specifications for designing accessible software. These specifications are also applicable to websites. ISO 9241-171 (2008) defines accessibility as the usability for people with the widest range of capabilities. The accessibility definition provided seems to extend the usability definition: "the extent to which a system, product or service can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use" (ISO 9241-210, 2010), to people with disabilities. This can apply as the basis for evaluating accessibility in terms of effectiveness, efficiency and satisfaction for people with the widest range of capabilities (Bevan, Carter, & Harker, 2015).

The BS 8878:2010 Web Accessibility Code of Practice (BSI, 2010), which provides guidance how to design accessible websites, defines accessibility by adopting definitions from the ISO and the WAI. The accessibility definition provided refers to usability for people with the widest range of capabilities, by referring to qualities of usability such as effectiveness, efficiency and satisfaction. The definition also refers to the different levels of interaction with the web content, for example, perceive, understand, navigate and interact with the web. The standard, which is mainly focused on the process of creating accessible websites, suggest the use of web accessibility guidelines to direct the development of websites. A set of different evaluation methods to assure the accessibility through the development of a website is also suggested, including conformance of websites to the guidelines, testing with assistive technologies and user-based studies with disabled users.

Standards seemed to influence the way web accessibility is evaluated. Many countries adopted standards as legislation for an accessible web. Several countries adopted WCAG 2.0 (e.g. Australia, Canada, United Kingdom) or a variation (e.g. France, Germany, China) standard to government agency websites or even commercial websites. The adoption of standards as regulation in many countries seems to influence the way accessibility is evaluated. The majority of the studies that can be found in the literature evaluate the accessibility of websites based on the conformance to guidelines, either WCAG or Section 508 (Baazeem & Al-Khalifa, 2015). Simply put, the need to comply with the laws and policies established in each country seems to have directed conformance to the guidelines as a standard method for measuring web accessibility.

2.1.2. Accessibility in relation to usability

Some of the standards refer to accessibility with usability properties, however, the relationship between accessibility and usability is still unclear. Different viewpoints of this relationship can be found in the literature (Petrie & Kheir, 2007; Shneiderman, 2002; Thatcher et al., 2002).

Shneiderman (2000, 2002) proposed the term "universal usability", that encompasses both usability and accessibility problems. That means that usability can be expanded in order to also address accessibility problems.

The definition of web accessibility by Thatcher et al. (2002) suggests that accessibility is a subset of usability, meaning that accessibility problems are particular types of usability problems. That implies that accessibility problems affect only disabled people without having any effect on non-disabled users. However, they also stated that usability problems affect all users equally, regardless of ability or disability.

Petrie and Kheir (2007) proposed a definition of web accessibility describing accessibility in terms of usability characteristics, such as effectiveness, efficiency and satisfaction. In order to understand the relationship between accessibility and usability, Petrie and Kheir conducted a study with six blind and six sighted users investigating the problems found by both user groups. The study showed that the problems encountered by both user groups could be grouped into three categories: problems encountered by blind users only (pure accessibility problems), problems encountered by sighted users only (pure usability problems) and problems encountered by both user groups (universal usability problems). The study showed that there are problems encountered by each user group separately, but there were also problems affecting both users group. This result shows that usability problems do not encompass accessibility problems (Shneiderman, 2000, 2002) and that accessibility problems are not a subset of usability problems (Thatcher et al., 2002).

A similar result was also found in a study conducted by Rømen and Svanæs (2012). The problems found by three blind, two physically impaired and two dyslexic users were compared with the problems found by six sighted non-disabled users. The study found a similar distribution of problems between disabled and non-disabled users with the study conducted by Petrie and Kheir (2007). There were problems encountered only by disabled users, problems encountered only by non-disabled users, and problems encountered by both, disabled and non-disabled users.

As can be seen from the literature accessibility and usability are related, as there are problems that are encountered by each user group separately, but there are also problems that may affect both user groups.

2.1.3. Accessibility in relation to user experience

The term user experience goes beyond the scope of usability and accessibility. User experience is a concept that starts with users and their relation to the technology and is often described with a variety of meanings ranging from usability to beauty and emotions of using a system. Hassenzahl and Tractinsky (2006) suggest that user experience encompasses pragmatic and hedonic attributes. Pragmatic attribute embodies qualities that are related to usability, for example, effectiveness and efficiency. Hedonic attribute includes qualities that are related to the user's emotional state. Morville (2005) proposes a framework that shows user experience being formed by seven facets, including accessibility and usability, which match with the (Hassenzahl & Tractinsky, 2006) model.

To understand blind users' experience on the web, previous research focused on the problems blind users encounter on the web and their performance (Disability Rights Commission, 2004; André Pimenta Freire, 2012; Power et al., 2012), the investigation of blind users navigation techniques and coping strategies (Bigham, Cavender, Brudvik, Wobbrock, & Ladner, 2007; Power et al., 2013; Theofanos & Redish, 2003; Vigo & Harper, 2013) as well as emotional aspects of users (Aizpurua et al., 2016; Lazar, Feng, & Allen, 2006).

Aizpurua et al. (2016) explored the relationship of web accessibility with different emotional user experience attributes. Web accessibility was measured by conformance of the website to the guidelines and the perceived accessibility of the website as rated by 11 blind users. For measuring user experience the researchers used the AttrakDiff 2 questionnaire that measures hedonic and pragmatic qualities of a website (Hassenzahl, 2004), and analyse the emotions of the participants using the emotion word prompt list by Petrie and Precious (2010), which consists of eleven emotional words that can be rated by the participants.

The study showed that perceived web accessibility is associated with most of the user experience attributes. A strong correlation was found between perceived accessibility and hedonic qualities, such as professional, classy, valuable, inclusive, bring me closer to people, and presentable. This indicates that participants felt closer to the websites that they experienced to be accessible. A correlation was also found between perceived accessibility and pragmatic qualities, such as simple, practical, clear, manageable, direct. Regarding the relation between perceived accessibility and emotional words, it was found that there was a positive relationship to words with positive meaning such as happy, interested, pleased; and negative relationship to words with negative meaning such as annoyed, bored, confused, disappointed, frustrated.

Regarding the accessibility of the website based on the conformance to the guidelines and participants' ratings on user experience measures, there was evidence to support the relationship between hedonic qualities of user experience and conformance to the guidelines only. It was found that websites that had higher conformance to the guidelines were perceived innovative, exciting and original, whereas websites that had lower conformance to the guidelines were perceived typical, conservative and lame.

Overall, the results of the study suggest that accessibility, as perceived by the users, can affect qualities of user's experience on the web as participants were feeling better in websites that were perceived accessible in comparison to websites that were perceived less accessible. Thus, accessibility can be considered a quality measure for creating a better user experience on the web for blind people. There was a few support to suggest that accessibility based on the conformance to the guidelines is related to user experience attributes. This result highlights the importance of considering accessibility as a quality of users' experience with the website rather than a property of a checklist that a website conforms to. Moreover, this result suggests that if we want to create a better accessible user experience for everyone the field needs to move accessibility from an assessment of technical accessibility standards, which is the most commonly used method to study the field (Baazeem & Al-Khalifa, 2015).

In order to create great user experiences for everyone, designers and developers should first understand the audience, which includes people with disabilities as well. By understanding the persons, it helps designers and developers to understand the diversity of accessibility needs (Horton & Quesenbery, 2014). Horton and Quesenbery (2014) provided a framework of principles to help designers and developers to create accessible user experiences on the web. The approach is based on not retrofitting accessibility as the last checklist of additions into the website, but incorporating accessibility through the design and development process of the website.

Horton and Sloan (2014) note that "accessible user experience brings the benefits of good user experience to people with disabilities" and that it can increase disabled users' satisfaction and enjoyment on the web. It was also pointed out that web accessibility approaches focused on the technical compliance to guidelines are not adequate to ensure a quality user experience (Sloan & Kelly, 2011). In order to create a quality user experience for disabled people, organisations should establish a practice that commits to accessibility that is adopted by every member of the development team.

2.1.4. Discussion for what is web accessibility

A range of definitions of web accessibility from standards and the literature have been presented. As can be seen, there is not a consensus on a single definition. In some cases, the definition relates to meeting the standards, which may indirectly relate to some user needs. In some other cases, it relates to qualities of usability, such as effectiveness, efficiency and satisfaction with disabled users.

The inability to reach a consensus for a definition of the concept can be problematic for the research community. If researchers do not have a common understanding of the concept, it is difficult to talk about web accessibility cohesively. Researchers cannot speak with certainty of what they are varying and controlling in their studies in relation to the concept. This makes research studies more difficult to compare, and it is not clear what knowledge is new. Without an agreement to a definition that encompasses all the components of the concept, researchers may not have a full picture of the impact of web accessibility on users' interaction on the web and may miss important components that may influence users' experience.

A unified definition that encompasses all the components of the concept is necessary, for researchers to better understand what the key components of the concept are. As a research community, we should have a concrete definition of web accessibility that can be used from the studies operationalising the concept.

To illustrate why this is problematic, consider what is happening within the field of web accessibility. Many researchers adopt the definitions proposed by the standards and study accessibility based on the conformance of the website to the guidelines, either WCAG (Gonçalves, Martins, Pereira, Oliveira, & Ferreira, 2012; Kuzma, 2010; Pribeanu, Gheorghe-Moisii, & Fogarassy-Neszly, 2015; Shi, 2007) or Section 508 (Lazar et al., 2011; Olalere & Lazar, 2011; Wentz et al., 2014). However, other pieces of work look at the impact they have on disabled users' experience on the web (Disability Rights Commission, 2004; André Pimenta Freire, 2012; Power et al., 2012). Other researchers adopt the definitions that relate accessibility to usability and study accessibility via usability attributes, such as effectiveness, efficiency and satisfaction, with disabled people (Coursaris, Swierenga, & Whitten, 2014; André Pimenta Freire, 2012; Lazar, Olalere, & Wentz, 2012). Given that the overlap between the user problems and the web accessibility guidelines is not a complete match, the two appear to be discussing slightly different concepts, which potentially leads to fracturing the assessment of the field of web accessibility.

Another limitation that can be observed from studies that operationalise the concept is that only a few of them report the definition of web accessibility that was adopted in their study (Abu-Doush, Bany-Mohammed, Ali, & Al-Betar, 2013; Brebner & Parkinson, 2006; Coursaris et al., 2014; André Pimenta Freire, 2012; Petrie & Kheir, 2007). Not reporting the definition used to operationalise web accessibility studies can be a key problem. As there is no clear or universally accepted definition of web accessibility, it is not clear what researchers are controlling and varying in relation to the definition. Thus, this makes it difficult to compare studies between them. Having one definition would lead to unification of rules that can be used to study web accessibility.

In Chapter 3, the definition of web accessibility will be explored more completely.

2.2. Problems blind users have on the web

A number of studies can be found in the literature that studied accessibility with blind users. Tables 56 and 57 in Appendix A. show the studies with blind users on the web that can be found in the literature.

2.2.1. Studies that look into the problems blind users have on the web

The first large study that can be found in the literature that looks into the problems blind users encounter on the web was conducted by Coyne and Nielsen (2001). The study comprised of two parts. During the first part of the study, an investigation into the problems that users with disabilities have on sixteen websites was conducted. Participants comprised of 35 visually impaired (18 blind and 17 partially sighted) and nine physically impaired users.

The second part of the study involved a comparison between disabled users' performance and the one of a sighted control group. Sixty participants comprised of 20 blind, 20 partially sighted and 20 sighted were asked to perform four tasks, three on specific websites and one on a non-specified website. An investigation of the participants' success rate, the time required to perform the task and the participants' perceived rating of satisfaction, confidence and frustration, was conducted. Regarding participants' tasks success rates, sighted participants had high success rates (78.2%), whereas the figures for disabled users were much lower, 12.5% for blind and 21.4% for partially sighted participants. Disabled users also required twice as much the time than sighted participants to perform the tasks. Blind participants required the longest time (16 minutes and 46 seconds), partially sighted participants required 15 minutes and 26 seconds, whereas sighted participants required 7 minutes and 14 seconds. In regard to the participants' subjective rating, the average response for their satisfaction, confidence and frustration was analysed (using a scale 1 to 7, with 7 being the most positive answer). Blind participants' mean subjective rating was 2.5, partially sighted participants' rating was 2.9, whereas sighted participants' mean rating was 4.6.

Based on the results of both parts of the study, the authors proposed a set of 75 web design guidelines. The authors provided the reasoning behind the proposed guidelines based on the problems that users encounter. By further looking into the reasoning of the guidelines, some of the problems that blind users encounter on the web can be extracted. This included images without alternative texts, images with non-descriptive alternative text, non-descriptive button titles, difficulties in skipping content on the page, difficulties in using the pop-ups, links opening new browser windows without any indication, too many links on the page, instructions using sensory characteristics, pages difficult to navigate, irrelevant content on the page, error messages conveyed through colour only, required form fields not clear, content not in appropriate order, content not descriptive, and use of tables for visual design instead of organising information.

This was one of the first large studies that increased our understanding of the problems blind, partially sighted and people with physical disabilities encounter on the web. In addition, the study pointed out that disabled users' experience on the web is at a disadvantage in comparison with sighted users. The presence of alternative research-based guidelines highlighted the drawbacks of WCAG as a standardised tool for creating accessible and usable websites from their early stages. However, a drawback of the proposed guidelines is that they are focused on specific user groups, rather than a diversity of disabilities.

In one of the biggest studies ever made studying web accessibility (Disability Rights Commission, 2004), an evaluation of 100 websites was performed with 50 disabled users. Participants comprised of blind, partially sighted, dyslexic, physically impaired and hearing impaired users. The blind participant panel comprised of ten blind screen reader users. Each participant was asked to evaluate ten websites by performing two tasks on each website. The results showed that blind participants had the worse task success rates amongst the other disabled user groups, with only 53% success rate. The difference between the five disabled user groups was also noted in the perceived rating of the task difficulty. The blind user group was the one with the lowest mean rating of task difficulty, meaning that blind users perceive the tasks more difficult amongst the other user groups.

To further understand the differences between blind users' and sighted users' experience on the web, the researchers evaluated six of the websites with a sighted control group. This further investigation compared the differences between blind users, the group that encountered most of the difficulties on the web, and a matched group of sighted users. The six websites selected comprised of three with high accessibility rating and three with low accessibility rating. On the sites with high accessibility rating, it was found that both user groups performed very well. On the sites with low accessibility rating, it was found that blind users completed only 67% of the tasks, whereas sighted users completed all the tasks. Regarding the time participants needed to perform the tasks, blind participants required around three to five times more time, depending on the accessibility rating of the website. Interestingly, on the websites with low accessibility rating, both user groups required a longer time to perform the tasks in comparison with the websites with higher accessibility rating.

The study reports the key problems found by blind users. This included problems with incompatibilities with the screen reader software and the web content, links and form label not being descriptive, no labels associated with input controls, cluttered and complex page structures, images without alternative text, and images with inadequate alternative text.

Disability Rights Commission (2004) study is the largest study that could be found in the literature that documents problems from five different disabled user groups. The study has gone some way towards enhancing our understanding of the experience of different user groups with disabilities on the web. In addition, the study shows that blind users are the user group that experienced most difficulties on the web. The study would benefit more if it included a comparison between the problems found between blind and sighted users. Moreover, the study suffers from a methodological drawback. The majority of the tasks were undertaken by users were unmoderated. Thus, users’ may not always be sure that they completed the tasks successfully and the problems encountered using the websites could not be observed by experts.

Petrie and Kheir (2007) examined participants' experience on the web and the relationship between the problems encountered on the web by blind and sighted users. Twelve participants, six blind and six sighted were asked to perform seven tasks on two mobile websites while performing a concurrent verbal protocol (CVP).

Regarding participants' efficiency, an analysis of participants' task rates was performed. The study showed that there was significant difference between the two users groups, with sighted users having higher task success rates in comparison with blind users.

With regards to the problems encountered, it was found that blind users encountered more problems than sighted users on both websites. Blind users encountered 288 problem instances, whereas sighted users encountered 192 problem instances on both websites.

A novel contribution of this study was the comparison of the problems found between the two user groups. The study demonstrated that there are problems that are encountered only by blind users (pure accessibility problems), problems encountered only by sighted users (pure usability problems) and problems encountered by both user groups (universal usability problems). Most of the problems reported in the study (62% of the problems) were pure accessibility problems. Pure usability problems comprised one-quarter of the total problems (24% of the problems). Universal usability problems were relatively low, with only 14% of the problems. A further analysis of the universal problems was conducted to investigate whether one user group perceived the problems more severe than the other user group. The analysis revealed that there was only significant difference on the severity ratings of the problems found between the two user users on one of the two websites. Blind users perceived those problems more severe than sighted users.

The study provided a better understanding of the relation of accessibility and usability problems, a subject that is under debate (see subsection 2.1.2). The importance of conducting studies with blind users was highlighted as there are problems that affect both user groups but there are many problems that only affect blind users. The study would have benefit if more participants were included in the study and if an analysis of the problem types that were shared and distinct to each user group was performed.

Rømen and Svanæs (2008, 2012) undertook a similar study with Petrie and Kheir (2007), with the difference of the inclusion of participants from more disabled user groups. An evaluation of two government websites was performed by seven disabled and six non-disabled users. Disabled participants comprised of three blind, two physically impaired and two dyslexic users.

Regarding the problem instances, it was found that disabled users experience significantly more problems than sighted users. Moreover, the study found a similar distribution of problems with Petrie and Kheir (2007). Pure accessibility problems were 59% of the problems, pure usability problems were 22% of the problems, and universal usability problems were 19% of the problems. The study also reports some of the problems that were distinct to each disabled user group. For blind users, this included problems with difficulties about links, such as links not being descriptive, too many links, duplicate links. The study also reports problems experienced by all user groups, for instance, lack of instructions on how to use the forms, or pop-ups not easy to use.

The study is subject to some limitations. First, the number of participants in each disabled user group was relatively low. In addition, as the way the three disabled user groups interact with the websites differ, a comparison between each disabled user group and sighted users regarding the number of problems and the distribution of problems they encountered would have been more interesting.

Stenitzer, Putzhuber, Nemecek, and Büchler (2008) evaluated five e-commerce websites with 14 participants. Participants comprised of older people, partially sighted and blind users. However, the number of participants of each user group was not specified by the authors. The study reports some of the main problems users encounter. This includes missing or inadequate labels for links and buttons, the position of elements not following users' expectations, important information not positioned at the top of the page, too much information on the page, navigation through the page content was difficult, disturbing advertisements, unclear labels. Although the study reported many of the problems users encountered, it presents the problems across all the three user groups. This makes it difficult to understand the problems encountered only by blind users.

Another large study regarding the number of participants involved (André Pimenta Freire, 2012; Power et al., 2012), investigated the problems encountered by three disabled user groups on the web. Participants comprised of 32 blind, 19 partially sighted and 12 dyslexic users. Participants evaluated 16 websites, by performing two or three tasks on the websites while performing CVP. The ISO 9241-171 (2008) definition of web accessibility was used to operationalise the study of the concept. Thus, the measures used were participants task success rates, perceived task difficulty ratings and number of problems encountered.

Regarding task success rates, blind participants were able to complete only 56% of the tasks, a similar figure to the task success rate found by Disability Rights Commission (2004). Regarding the difference between the three user groups, it was found that both blind and partially sighted had a lower percentage of succeeded tasks than dyslexic users, but had no difference between them. Regarding the perceived task difficulty, both blind and partially sighted users perceived the tasks as more difficult than dyslexic users. With regards to the number of problems encountered, blind users encountered most of the problems reported (1383), whereas partially sighted users encountered 936 and dyslexic users 693 problems. There was only significant difference on the number of problems encountered between either blind or partially sighted, and dyslexic participants.

The study reports a list of the most frequent problems encountered by blind users. This includes links' description not being clear, content not found in pages where users might have expected it, irrelevant content on the page, form controls not clear what will do, no or insufficient feedback on user's actions, confusing content, page functionality not working, no headings, elements not reachable via screen reader, inadequate alternative text on images and no enhancement to audio, video or multimedia.

The distribution of the problems between the three user groups was also analysed. The overlap of problems between the three user groups was relatively low, counting 4.1% of the problems. The overlap of problems between blind and partially sighted was 4.7% of the problems, whereas the overlap between blind and dyslexic users was 2.8% of the problems. The overlap of the problems between dyslexic and partially sighted was 5.5% of the problems. The small figures of overlap between the three disabled user groups demonstrated that there are problems distinct to each user group. This highlights the importance of conducting studies with more than one disabled user groups in order to address the needs of diverse users.

The study also looked further into the problems that were distinct to each disabled user group. For blind users, this included problems with the headings of the page, images without alternative text, too many or duplicate links, tables not well structured.

The study provided a large corpus of problems encountered by blind, partially sighted and dyslexic users. As the aim of the study was to investigate the coverage of the problems found by the accessibility guidelines, the study did not include a control sighted user group.

A study that also included a large number of blind participants was conducted by Lazar et al. (2012). An investigation of the accessibility of 16 job seeking websites in eight states of the USA was conducted. Sixteen blind participants were asked to perform two tasks on the websites while performing CVP.

Participants task success rate was relatively low, with only 28.1% of the tasks being completed without any assistance. Some of the problems blind participants encountered on the websites involved incompatibilities between the assistive technology and page content, confusing instructions, error message not helping users recover from their errors, no labels associated with interactive elements, required fields on forms not specified, areas inaccessible via screen reader, no use of bypass links, content not presented in a logical tab order, input formats not clear, tables not well structured.

The study provides a descriptive list of problems found across the 16 websites. However, the study is subject to some limitations. From the description of the study design, "two attempts were made to apply for jobs on each Web site (for a total of 32 attempts at submitting a job application)", it seems that each website was evaluated by either one participant performing two tasks or by two participants, each performing a task.

Coursaris et al. (2014) investigate the accessibility and usability of a health website in the USA, using the definition by ISO 9241. Twenty-five participants, comprised of 16 blind, four partially sighted and five sighted. Participants undertook seven tasks on the website while performing CVP. Regarding participants' task success rates and time to perform the tasks, the authors report that in most cases users with visual disabilities had lower success rates and required more time in comparison to sighted users. However, it was not statistically tested. With regards to the problems participants encountered on the website, the study organises the issues based on the task. Some of the issues encountered by blind users included content not being clear, pop-ups not accessible, no feedback on user's actions, functionality not as expected, abbreviations not explained, difficulties using input elements, no labels associated with interactive elements. Even though the study included three different user groups, no direct comparison between the user groups was conducted regarding task success rates and time, and the differences regarding the type of problems encountered between the three user groups.

2.2.2. Common problems for blind users

A review of the studies that can be found in the literature with blind users shows some commonalities on the problems reported. Table 57 in Appendix A. list the problem types reported in studies with blind users on the web. Some of the common problems reported by blind users across the different studies are no labels associated with input controls (R. Babu, 2013; Brebner & Parkinson, 2006; Coursaris et al., 2014; Disability Rights Commission, 2004; André Pimenta Freire, 2012; Lazar et al., 2012), difficulties navigating through the page because of complicated page structures (Craven, 2003; Disability Rights Commission, 2004; Lazar et al., 2012; Ramayah, Jaafar, & Yatim, 2013; Yoon, Newberry, Hulscher, & Dols, 2013), non-descriptive links (Abu-Doush et al., 2013; R. Babu & Singh, 2013a; Brebner & Parkinson, 2006; Byerley & Beth Chambers, 2002; Disability Rights Commission, 2004; Federici et al., 2005; André Pimenta Freire, 2012; Lazar et al., 2012; Oppenheim & Selby, 1999; Rømen & Svanæs, 2012), images without alternative text or images with inadequate alternative text (Abu-Doush et al., 2013; Brebner & Parkinson, 2006; Coyne & Nielsen, 2001; Disability Rights Commission, 2004; Oppenheim & Selby, 1999).

Research trying to address some of the problems

Some researchers investigated different design solutions to some of these issues. For example, T. Watanabe (2009) looked the benefits of structuring the page content using headings on blind users' experience. Two websites, one structured with headings and one without headings but with the same appearance were used. Sixteen sighted and four blind participants took part in the study. Participants were asked to perform four tasks on each of the two websites. The study showed that both user groups required less time to perform the tasks on structured websites, with the benefits of task time be greater for blind users.

Power, Petrie, Freire, and Swallow (2011) compared the effectiveness of different techniques for providing descriptive links. Eleven techniques from WCAG 2.0 (Reid & Snow-Weaver, 2008) were evaluated with 22 blind and three partially sighted participants. Participants were asked to undertake a specific task on the websites and to identify where the link would lead them. The study found that only one of the techniques, describing the link purpose in the text of the link, helped participants to correctly determine the link target destination almost all of the times. The study provided empirical results of the benefits of specific techniques for implementing links on pages and suggests which are the best approaches for providing links that are accessible and users can have high confidence regarding their destination.

These two studies provided empirical evidence of how specific design solutions can benefit blind users' experience on the web. In order to provide solutions to the problems, we must first develop a deep understanding of the problems. Then, solutions can be suggested and tested how and if they benefit users' experience.

2.2.3. Limitations of studies that look into the problems blind users have on the web

What we know about the problems blind users have on the web is based on studies that investigate the accessibility of websites with blind users. However, most of the previously mentioned studies suffer from some serious limitations.

Composition of participant pool

A limitation of the studies with blind users is the low number of users. The majority of the studies had a very low number of users, with a few exceptions (Coyne & Nielsen, 2001; Disability Rights Commission, 2004; André Pimenta Freire, 2012; Lazar et al., 2012). Half of the studies had six or less blind participants. There were even studies with only one blind participant (Oppenheim & Selby, 1999; Ramayah et al., 2013). Nevertheless, there are studies that the exact number of blind participants is unclear. For instance (Federici et al., 2005) reported that there were "four students with visual disability (one blind students and three with diminished vision)", without reporting any further information about the participants with diminished vision.

Jaeger (2006) studied the accessibility of ten government websites in the USA with ten participants, "ten participants had either a visual impairment or mobility impairment ... the users with visual impairments included individuals with no vision, low vision, double vision, and inability to focus". The number of blind screen reader participants is not explicitly reported by the author. The study reports a list of the user problems found on each website. However, it is not clear which of those problems were encounter by blind users.

Stenitzer et al. (2008) reported “14 user tests have been conducted so far within the two target groups: older adults between 50 and 72 years and people with visual impairment and blind users.”, without reporting any further information about the number of participants in each user group. Moreover, the problems users encountered are not reported by user group, which makes it difficult to understand which of those problems were encounter by blind users.

As can be seen from the review of the literature with studies with blind users, the number of participants included in studies is relatively low, with the exception of just a few studies. Moreover, there are studies that their results cannot be taken into consideration as there is a serious weakness in their methodology, as they do not report important information (Federici et al., 2005; Jaeger, 2006; Stenitzer et al., 2008).

Comparison with sighted users

In reviewing the literature there are studies that included both blind and sighted users (Coursaris et al., 2014; Coyne & Nielsen, 2001; Craven, 2003; Disability Rights Commission, 2004; Giraud, Colombi, Russo, & Thérouanne, 2011; Petrie & Kheir, 2007; Rømen & Svanæs, 2008, 2012; Swierenga, Sung, Pierce, & Propst, 2011; Yoon et al., 2013). However, not all studies made a comparison between the two user groups (Swierenga et al., 2011; Yoon et al., 2013).

Comparing the experiences between blind and sighted users it can provide a further understanding of the differences between the two user groups in regards to the problems encountered on the web and their experiences. Thus, design solutions that properly address the issues can be proposed and evaluated. However, solutions cannot be suggested unless the problems are fully analysed, and this involves a thorough understanding of the causes.

Swierenga et al. (2011) investigated the accessibility of an information website. Participants comprised of eight blind, eight partially sighted and 18 sighted participants. The study provides results regarding participants' time to perform the tasks and task completion rates. Regarding participants task completion, the study reports that participants were more successful in pages without complex tables.

However, the study does not specifically report any individual differences between the three user groups regarding task completion. As the three user groups interact with the websites differently, presumably they might be differences in regards participants difficulties in completing the tasks. For participants' task completion time, the study reports that blind participants required more time than both partially sighted and sighted participants, although it was not statistically investigated.

Yoon et al. (2013) conducted a study with blind and sighted participants investigating the accessibility of five websites, three library websites and two non-library websites. Participants comprised of six blind screen reader users. The authors also report: "usability tests were also conducted with sighted users for comparison purposes". However, the study does not report the number of sighted users. In addition, the study reports some preliminary results from the analysis. The authors did not report any results regarding the comparison between blind and sighted users.

There are also studies that included other disabled user groups as well, but reported the results from the disabled user groups unified (Craven, 2003; Rømen & Svanæs, 2008, 2012).

Very little was found in the literature about the differences between blind and sighted users' experience on the web (Coursaris et al., 2014; Coyne & Nielsen, 2001; Disability Rights Commission, 2004; Giraud et al., 2011; Petrie & Kheir, 2007). However, none of these studies attempted to address the question whether the type of problems encountered by blind users was specific to blind users or problems that everyone encountered. Petrie and Kheir (2007) demonstrated that there are problems distinct to each user group but also problems that are encountered by both user groups, yet no research was conducted to further look into the problem similarities and differences.

Limitations on the method used to identify the problems

In user-based studies of websites, typically a number of users perform a number of tasks on the target website. The most basic user evaluation involves users performing the tasks in order to measure the users' performance on it. In addition, users can perform the tasks while performing a verbal protocol.

Table 56 in Appendix A. list the verbal protocol used in the studies with blind users on the web. As can be seen, most of the studies used a verbal protocol in their study. The verbal protocol can be an effective tool in the hands of evaluators as it can offer insight into the users' thought process, their problem-solving strategies (Nielsen, 1994) and it can be an effective method for detecting the problems users encounter with a system (Jørgensen, 1990; Wright & Monk, 1991). The verbal protocol was first introduced in human-computer interaction studies by Lewis (1982), but its origins can be traced back to the work of Ericsson and Simon (1980, 1993), in cognitive psychology. Many usability textbooks have established the verbal protocol as a core component of usability testing practice (Dumas & Redish, 1999; Nielsen, 1994; Rogers, Sharp, & Preece, 2007; Rubin, 1994).

Different approaches (Boren & Ramey, 2000; Dumas & Redish, 1999; Ericsson & Simon, 1980, 1993; Olmsted-Hawala, Murphy, Hawala, & Ashenfelter, 2010) and methods can be found in the literature on how to perform the verbal protocol. There are approaches were the only interaction between the participants and the evaluator is to remind them to think aloud (Boren & Ramey, 2000; Ericsson & Simon, 1980, 1993), but also approaches were the evaluator has more active role and can ask directly questions the participant about different areas of the website (Dumas & Redish, 1999; Olmsted-Hawala et al., 2010). Regarding the method, the verbal protocol can be performed either concurrently, concurrent verbal protocol (CVP), or retrospectively, retrospective verbal protocol (RVP). In CVP participants think out loud while doing the task, whereas in the RVP participants first perform the tasks in silence and then they perform the verbal protocol, usually prompted by a video of themselves performing the tasks (Nielsen, 1994; Rogers et al., 2007). Blind participants can also perform RVP by listening to an audio of their interaction with their screen reader. The choice of the verbal protocol approach and method may have an impact on the results of the study.

Russo, Johnson, and Stephens (1989) argued that performing CVP can be reactive, meaning that it can improve participants task performance, which may lead to failure of detecting user problems or worsen user's task performance, which may lead to identifying false positive user problems. Studies investigated the reactivity of CVP with different approaches and found evidence to support the claims of Russo et al. (1989). However, reactivity is mainly dependent on the selected approach used. Studies demonstrated that either Ericsson and Simon (1980, 1993) or Boren and Ramey (2000) could be considered an appropriate approach to use as neither of them is reactive (Bruun & Stage, 2015; Olmsted-Hawala et al., 2010).

Studies investigated whether there are differences between the two verbal protocol methods, CVP and RVP. Based upon the results of previous research (Alshammari, Alhadreti, & Mayhew, 2015; Bowers & Snyder, 1990; Van den Haak, De Jong, & Jan Schellens, 2004, 2007,

2009) the two verbal protocols are comparable regarding participants' task success rates and time needed to perform the tasks. However, regarding the number of problems the two protocols reveal, the results are still unclear. Some studies suggest that RVP can identify more problems reported from users (Van den Haak, De Jong, & Jan Schellens, 2003; Van den Haak et al., 2004), while others suggest that the two protocols are comparable (Van den Haak et al., 2007, 2009) and others suggest that CVP identifies more problems (Alshammari et al., 2015). Van den Haak et al. argue that the difference of the problems found between the two protocols may be the result of the double workload of the participants in the CVP condition, having to perform the tasks and think aloud simultaneously. In RVP, participants need to perform only one task, which can provide them with the opportunity to report more problems. Hertzum, Hansen, and Andersen (2009) found evidence to support the claim that performing CVP has a higher workload than not performing a verbal protocol. It was found that CVP was more mentally demanding than doing the tasks in silence.

Even though a number of studies investigated the differences between CVP and RVP, some of the studies are subject to some methodological limitations. Alshammari et al. (2015) did not state how the problems identified, whether they were reported by the users or found by experts. Regarding Van den Haak et al. studies, participants in all of the studies were students between the ages of 18 and 24. Thus, the results of the studies cannot be generalized to all people.

While there is some evidence that there are differences between the two verbal protocol methods, it seems that CVP is the preferred verbal protocol method among researchers and practitioners in the field of user testing (McDonald, Edwards, & Zhao, 2012). The main reason why researchers and practitioners prefer CVP is because it is fast to use. On the contrary, some respondents noted that RVP demands more time and they may not often have time to review the videos as their resources are limited.

A similar preference of verbal protocol on the studies conducted with blind users can be observed from Table 56 in Appendix A. Even though CVP seems the norm in user-based studies with blind users there are potential impacts and consequences of this choice, and researchers may not be able to capture the full picture of the problems blind users encounter on the web. It was suggested that other approaches should be considered when conducting user-based studies with blind users (Chandrashekar et al., 2006; Strain, Shaikh, & Boardman, 2007). Some studies pointed out that blind participants did not often respond to prompts to continue verbalising their thoughts when performing CVP (Chandrashekar et al., 2006; Coyne & Nielsen, 2001). It was also noted that blind users might have a difficult time to perform CVP, as they would have to verbalise their thoughts and listen to the screen reader output the same time (Strain et al., 2007). If the CVP may not be the most appropriate method to use with blind users, other verbal protocol methods should be considered. It has been suggested that RVP can be considered a better option for studies with blind users (Chandrashekar et al., 2006; Strain et al., 2007). However, no research could be found investigating the two verbal protocols with blind users.

Discussion of research on problems of blind users on the web

Previous research looked into the problems blind users have on the web. However, our current understanding of the type of problems is not very clear. In addition, it is not clear whether the type of problems encountered by blind users are distinct to blind users or problems everyone has on the web.

Many of the studies that investigated the problems that blind users encounter on the web are subject to methodological limitations. It is important to grow the body of literature on the problems that blind users have on the web through research that will properly study the concept. Thus, our current understanding of the problems that blind users encounter on the web will be enhanced. By providing additional evidence in respect to the problems that are distinct to blind users, there can be practical applications, as developers and designers can design solutions that properly address these problems and test their benefits on blind users' experience. In addition, designers will be able to prevent these problems before they are embedded into the websites. It will also provide additional knowledge into prioritising the problems that first need to be addressed in existing website. Moreover, by distinguishing the problem types that are distinct to each user group, designers and developers can suggest and test solutions to the problems without changing the experience for other user groups. However, we must first develop a thorough understanding of the problems blind users have before design solutions can be suggested and tested.

2.2.4. Summary

The literature review reported in this chapter presents an overview of the different definitions of web accessibility and what we know about it. It shows that different definitions of web accessibility can be found and highlights the potential consequences of the lack of an agreement on a universally accepted definition of the concept. Then, it directs the attention to what we know about the problems blind users have on the web, by presenting studies that look into the problems. It provides an overview of common problems that reported across different studies with blind users, but also raise the limitations of the studies conducted with blind users. Moreover, it highlights the lack of previous studies of not investigating the differences between the problem types encounter on the web between blind and sighted users. Based on the results of previous research (Petrie & Kheir, 2007), it has been demonstrated that there are problems distinct to each user group, yet no research was conducted to further understand the problem type differences. It is important to understand the problem differences between blind and sighted users, in order to be able to propose design solutions to prevalent problems and test how and if benefit blind users' experience. However, solutions cannot be suggested unless a comprehensive understanding of the problem similarities and differences between blind and sighted users is provided.

Before conduct studies that investigate user problems and their potential solutions, there is an important gap in the literature that hinder the experimental research in this area. There is lack of agreement on a definition of web accessibility, which makes it difficult to operationalise experiments. Thus, the first study of this thesis will focus on what is web accessibility and what are the key components that researchers consider to be as part of the concept.

Chapter 3. What is web accessibility

3.1. Introduction

As can be seen from the review of the literature, different definitions of web accessibility with different viewpoints can found. This can be problematic for the research community as researchers do not have a common understanding of the concept. Also, it is difficult to design studies that properly study web accessibility as it is not clear what to vary and control in relation to the concept.

In order to understand the problems blind users have on the web compared to sighted users and design and test solutions to key problems, a series of studies that operationalise the concept needs to be conducted. However, first there is a need to have a clear understanding of the key components of the concept considered by researchers as part of web accessibility. Thus, the first study of this thesis will focus on identifying the core components of the concept based on the different definitions of web accessibility that can be found in the literature; and propose a unified definition of web accessibility that encompasses all the core components. The definition can then be used by the studies operationalising the concept.

The study addresses the following research question:

- What researchers consider to be the key components of web accessibility?

3.2. Method

In order to create a unified definition of web accessibility that can be used in future studies, a content analysis was undertaken of existing definitions of web accessibility that were found in the literature. Content analysis is a method that can be used to analyse document or elements of text to attain a greater understanding of particular phenomena through the concepts that emerge from the analysis (Elo & Kyngäs, 2008; Krippendorff, 2004).

The content analysis process followed is consisted of three main phases: preparation, organizing and reporting (Elo & Kyngäs, 2008). Preparation phase involves the unit of analysis and the sample selection. Organization phase involves the coding of data and creating higher order concepts. The reporting phase includes the description of the results of the analysis process.

3.2.1. Preparation

The preparation phase starts with the selection of the unit of analysis. In this study, a sentence or phrase mentioning particular concepts was considered as a unit of analysis. The number of definitions mentioning particular concepts was considered as measures.

Then, a selection of a representative sample of definitions of web accessibility was performed. Web accessibility is a concept that applies to various domains. For example, studies of web accessibility of government, tourism, health, libraries' websites can be found in the literature. In order to investigate the different definitions of web accessibility introduced, a comprehensive web search of publications on web accessibility was conducted. Nine digital libraries were searched for publications. This included ACM Digital Library, IEEE Xplore, Web of

Science, Collection of Computer Science Bibliographies, Scopus, Science Direct, Emerald Insight, Wiley and Google Scholar. The libraries were selected in order to cover a wide range of different sectors that could include publications on web accessibility.

Search terms used were "website/web site/web-site/webpage/web page/web- page accessibility" and "accessibility of website/web site/web-site/webpage/web page/web-page". The search terms selected were as broad as possible to capture as many relevant publications as possible. The search term "web accessibility" was not included as it was too general. For example, in Google Scholar it returned 68500 results, whereas in ACM Digital Library and Computer Science Bibliographies 1956 and 3005 results were returned respectively. In addition, many of the returned results were not relevant.

The search scope included the use of the exact phrase in the entire document, without any predefined year restriction. A limit was placed on the language of the document considering only documents written in English as relevant.

The pool of returned papers was then searched for possible introductions of definitions of web accessibility. This was achieved by searching for terms as "web accessibility is...", "web accessibility means...", "web accessibility aims...", "definition of web accessibility..." and by manually skimming the Introduction and Background section of the papers.

In total 157 definitions of web accessibility was found, including definitions from guidelines and standards, published papers, books and online documents. The definitions represent an international view of web accessibility, as they include definitions from authors from 35 different countries from all continents. The list of definitions can be found in Appendix B.

The search corpus was also used in another research project, not included in the thesis, in collaboration with other researchers. That involved an analysis of the methods used to study web accessibility and the state of web accessibility based on conformance evaluation to guidelines through the years.

3.2.2. Organizing

The organizing phase was conducted in two parts. First, all relevant content words were extracted from the definitions. Grammatical variations such as *understand* and *understandable* were grouped together. Similarly close synonyms (for example *anyone* and *all users)* were grouped together. The second part of the phase involved grouping the content words into higher order concepts. The aim of the grouping was to reduce the number of content words that are similar into higher level concepts.

3.2.3. Reporting

The last phase of the process involved a reliability check of the analysis and the presentation of the results from the analysis.

An inter-coder reliability was performed for a sample of 50 definitions. One of the two supervisors of the thesis, coded independently the definitions with satisfactory levels of agreement, with Cohen’s Kappa greater than 0.8 (Landis & Koch, 1977).

From the analysis, six concepts emerged. The concepts included:

- Groups of users, characteristics, needs of users
- What users should be able to do

- Technologies used
- Characteristics of the website
- Characteristics of the situations of use
- Design and development of the website

3.3. Results

Table 2 shows the number of definitions that included the six different concepts. The table also includes the number of important specific examples within the concepts with their frequencies.

Table 2. Concepts used in the 157 definitions of web accessibility.

Concept	**Explanations, examples with frequencies**	**Number of Definitions mentioning**
Groups of users, characteristics, needs of users	with disabilities (106), all users/as many as possible (95), characteristics (57), specific disabled groups (29)	156 (99.4%)
What users should be able to do	access (96), use (58), interact (36), understand (35), perceive (25), navigate (24), contribute (13), available (10), get (3), achieve goals/reach (2), benefit/perform/visit (1)	148 (94.3%)

Technologies used	mainstream technologies (39), assistive technologies (28)	61 (38.9%)
Characteristics of the website	usability or aspects of usability (efficiency, effectiveness etc.)	46 (29.3%)
Characteristics of the situations of use	in specified contexts of use (20), in environmental constraints (20)	39 (24.8%)
Design and development of the website	design (32), standards/guidelines (6)	37 (23.6%)

As can be seen from the Table 2 above, almost all the definitions referred to groups of users or the characteristics of users. However, a slight vagueness can be observed between the definitions. Many definitions refer to all users or as many users as possible (*N* = 95), whereas some others refer specifically to people with disabilities (*N* = 106) or to specific disabled groups (*N* = 29). It is interesting that some definitions explicitly refer to people with disabilities (see Table 3), whereas some others are more vague using terms like “anyone” (Letourneau, 1998), “all users” (Waddell, 1998), “all kinds of people” (M. Watanabe, Asai, & Asano, 2007), “as many people as possible” (Wang, Liu, & Hua, 2010). While people with disabilities are undoubtedly included in these terms, it is interesting that some definitions were more explicit than others by specifically referring to people with disabilities. Interestingly, only 11 definitions explicitly referred to older users (see Table 3). However, this user group is often considered in the scope of web accessibility and constitutes an important and growing proportion of the population. While these users are undoubtedly included in “all users” or “anyone” related references, it is odd that it appeared in so few definitions. Presumably, the user group is not often reported as many researched may not realise it might be a relevant group or forget to consider it when working on web accessibility.

Table 3. Examples of definitions that refer to groups of users or the characteristics of users.

All users/anyone/as many users as possible	
Letourneau (1998)	... *anyone* using any kind of web browsing technology must be able to visit any site and get a full and complete understanding of the information as well as have the full and complete ability to interact with the site if that is necessary.
Waddell (1998)	...the design of a webpage ... in order to ensure that *all users* can access the information on the page.
M. Watanabe et al. (2007)	Web accessibility means the ability [of websites] to be accessed *by all kinds of people* or devices.
Wang et al. (2010)	… the degree to which a product is *accessible by as many people as possible*…
People with disabilities	
WAI (2005)	... *people with disabilities* can use the Web. More specifically, Web accessibility means that people with disabilities can perceive, understand, navigate, and interact with the Web, and that they can contribute to the Web.
Paciello (1996c)	... the ability for [web] browsers to render information in a manner that is accessible to *people with disabilities*. For the blind, any aspect of a graphic interface presents barriers. For low vision web surfers (and in some cases, those with cognitive limitations), data presentation in different formats, different

	fonts, and inconsistent character and word spacing, make reading online information difficult. For the deaf, rendering sounds or sound bytes presents significant challenges.
Petrie and Kheir (2007)	the ultimate criteria for accessibility should be user-based and we can adapt the ISO 9241 definition for this purpose: the extent to which a product/website can be used *by specified users with specified disabilities* to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use.
Henry (2007)	…means that *people with disabilities* can use a product.
Specific disabled groups	
De Lima, Lima, and De Oliveira (2009)	Web accessibility is the degree to which *people with visual, auditory, physical, speech, cognitive, or neurological disabilities* can perceive, understand, navigate, and interact with the Web.
Curl and Bowers (2009)	Website accessibility can be defined as the ability to access the web regardless of "*visual, hearing, mobility or learning disabilities*", speed of Internet connection/bandwidth, or age of computer/software technology.
Older users	
Maswera, Dawson, and Edwards (2005)	... the most important component in web accessibility is addressing issues relevant to individuals with disabilities and the *elderly*.
Alexandru and Alecu (2010)	Web accessibility refers to making the World Wide Web accessible and available to everyone, including people with disabilities and *senior citizens*.
Park, Lim, and Lim (2014)	Web accessibility means ensuring that anyone including those with disabilities and the *elderly* can access all information provided by websites in any technical environment without much special skill.

Another interesting result of the analysis is the wide number of terms used to describe what people should be able to do as a result of web accessibility. A large number of definitions define web accessibility in terms of user actions, such as access (N = 96), use (N = 58), interact (N = 36) and understand (N = 35). It is interesting that the most frequently mentioned term that used to describe what people should be able to do is "access", which is a derivative of the portmanteau word accessibility, which combines the meanings of the words access and ability. Table 4 shows the variation of the different terms used to describe what people should be able to do as a result of web accessibility.

Table 4. Examples of definitions that used terms to describe what people should be able to do as a result of web accessibility.

Access	
Waddell (1998)	... the design of a webpage ... in order to ensure that all users can *access* the information on the page.
S. K. Kane (2007)	Web accessibility refers to the degree to which a website may be *accessed* by people with varying abilities.
Luján-Mora, Navarrete, and Peñafiel (2014)	… refers to creating websites accessible to all users who want to access them, regardless of users' disability. When websites are correctly designed and developed, all users can have *access* to their information and functionality... the objective of the web accessibility is to ensure that people with disabilities can *access* websites just like everyone else.
Use	
Section 508 (1996)	Technology is accessible if it can be *used* as effectively by people with disabilities as by those without.

Lazar, Schroeder-Thomas, et al. (2003)	An accessible web site is a web site that can be successfully *used* by people with various disabilities. People with different disabilities may be using different forms of assistive technology, such as screen readers, alternative keyboards, or alternative pointing devices.
Thatcher et al. (2002)	... people being able to get and *use* web content. It is about designing web pages that people can present and interact with according to their needs and preferences …
Interact and/or understand	
Akhter, Buzzi, Buzzi, and Leporini (2009)	… An accessible (Web) user interface means that potential technical barriers have been eliminated, and thus anyone can *interact* with it ...
Batra (2009)	The main goal of Web Accessibility is to make it possible for everyone to use, *understand* and communicate using Web based resources, despite any disabilities or technological restrictions.
WAI (2005)	... people with disabilities can use the Web. More specifically, Web accessibility means that people with disabilities can perceive, *understand*, navigate, and *interact* with the Web, and that they can contribute to the Web.
Yates (2005)	"... web accessibility is the ability for a person using any agent (software or hardware that retrieves and renders web content) to *understand* and fully *interact* with a web site's content.

One-third of the definitions refer to the technologies people are using. First, it is interesting that only one-third of the definitions refer to the technology people are using in relation to web accessibility. Technology is an important component when it comes to users' interaction with the website and it may impact their experience if it is behaving in an unexpected way. Second, some definitions were vaguer than others. Some definitions refer to the technology without specifying assistive technologies, while some others explicitly referred to examples of assistive technologies. Table 5 shows some examples of definitions that refer to the technology people are using. As can be seen from the table below, some definitions explicitly refer to assistive technologies in general (Fogli, Colosio, & Sacco, 2010; Zeng, 2004), whereas others refer to examples of assistive technologies (Jaeger, 2006; Lazar, Beere, Greenidge, & Nagappa, 2003; Maatta Smith, 2014; Trewin, Cragun, Swart, Brezin, & Richards, 2010), such as screen readers, screen magnifications, speech input, alternative keyboards. There were though definitions that refer to the technology people are using with terms like “any user agent” (Sierkowski, 2002) or “any kind of web browsing technology” (Letourneau, 1998). Certainly, assistive technologies are included in these terms. However, not an explicit reference to them has been made.

Table 5. Examples of definitions that refer to users’ technology.

Technology people are using	
Sierkowski (2002)	Web accessibility is the ability for a person using any <u>*user agent (software or hardware that retrieves and renders web content)*</u> to understand and fully interact with a website’s content …
Letourneau (1998)	... anyone using any <u>*kind of web browsing technology*</u> must be able to visit any site and get a full and complete understanding of the information as well as have the full and complete ability to interact with the site if that is necessary.
Assistive technology	
Fogli et al. (2010)	Website accessibility is mainly concerned with easy web content fruition by different categories of people, including

	those navigating the web through *assistive technologies*, which provide their users with alternative ways of accessing web pages.
Zeng (2004)	… Web accessibility can be defined as the degree to which it is accessible through *assistive technologies* used by persons with disabilities. (Zeng, 2004)
Specifying assistive technologies	
Lazar, Schroeder-Thomas, et al. (2003)	An accessible web site is a web site that can be successfully used by people with various disabilities. People with different disabilities may be using different forms of assistive technology, such as *screen readers, alternative keyboards, or alternative pointing devices*. A web site that is accessible is flexible enough to work with these various assistive technology devices.
Jaeger (2006)	For a web site to be accessible, it should provide equal or equivalent access to all users, and it should work compatibly with *assistive technologies such as narrators, screen enlargement,* and many other devices that persons with disabilities may employ to navigate cyberspace.
Maatta Smith (2014)	… accessibility refers to the viability of an individual with disabilities to access and use information as it is presented on the public library's website. Accessibility considers whether information can be read by manipulating text on the screen (enlarging text size, changing color and contrast) or *through the use of other adaptive technologies, such as screen readers or refreshable braille displays*.
Trewin et al. (2010)	... a Web application or page is accessible if people with disabilities - including people requiring *assistive technologies such as screen readers, screen magnifiers, or speech input* - are able to access any information from it and perform any operations it implements.

Characteristics of the website were also mentioned by almost one-third of the definitions. This concept mainly involves qualities of usability, such as effectiveness, efficiency, satisfaction, ease of use. Looking into some of the definitions that refer to the characteristics of the website (see Table 6), it can be seen that some definitions directly referred to usability as a broader concept, whereas some others refer directly to qualities of usability (De Oliveira Junior, Motti, Freire, & De Mattos Fortes, 2007; Petrie & Kheir, 2007), such as effectiveness, efficiency and satisfaction.

Table 6. Examples of definitions that refer to characteristics of the websites (e.g. usability or aspects of usability).

Usability	
ISO 9241-171 (2008)	... the *usability* of a product, service, environment or facility by people with the widest range of capabilities.
Chevalier, Dommes, and Martins (2013)	Web accessibility is the inclusive practice of making websites *usable* by people of all abilities and disabilities.
Qualities of usability	
Petrie and Kheir (2007)	... the ultimate criteria for accessibility should be user-based and we can adapt the ISO 9241 definition for this purpose: the extent to which a product/website can be used by specified users with specified disabilities to achieve specified goals with *effectiveness*, *efficiency* and *satisfaction* in a specified context of use.
De Oliveira Junior et al. (2007)	Accessibility is related to make a system *usable*, *efficient*, *effective* and to *satisfy* “more people in more different situations” …

There was, however, a slight discrepancy between a few definitions regarding the usability concept. Four definitions describe accessibility and usability as two completely distinct concepts. For example, Buzzi, Buzzi, Leporini, and Akhter (2009a) notes: “Accessibility is a basic pre-requisite for allowing users to have access to the web page content, while usability provides online users with simple, efficient, rapid and satisfying navigation and interaction” and Mori, Buzzi, Buzzi, Leporini, and Penichet (2011) notes: “Accessibility is a prerequisite that permits users to perceive online content and interact, while usability enhances the quality of the interaction, which should be simple, efficient and satisfying”. These definitions describe accessibility as a requirement to access the web content, whereas usability as a concept that improves users’ experience.

One-quarter of the definitions refer to the characteristics of the situation of use. For example, specified contexts of use or environmental constraints (see Table 7). Some definitions describe specific constraints, such as language or technological constraints (A. P. Freire, C. M. Russo, & R. P. Fortes, 2008). Also, almost a quarter of the definitions referred to the practice of creating accessible websites (see Table 7), of which very few mentioned the standards or guidelines (Craven & Nietzio, 2007). This is quite interesting, as even in the definitions of web accessibility that refer to the practice of making websites very few of them refer to the standards or guidelines.

Table 7. Examples of definitions that refer to specific context of use or environmental constraints, or the practice of making websites.

Specific context of use or environmental constraints	
Petrie and Kheir (2007)	... the ultimate criteria for accessibility should be user-based and we can adapt the ISO 9241 definition for this purpose: the extent to which a product/website can be used by specified users with specified disabilities to achieve specified goals with effectiveness, efficiency and satisfaction in a *specified context of use*.
Mankoff, Fait, and Tran (2005)	Web accessibility involves making web content available to all individuals, regardless of any disabilities or *environmental constraints* they experience.
Andre P Freire et al. (2008).	...Web accessibility corresponds to making possible to any user, using any user agent (software or hardware to view Web content) to understand and interact with a Web site, despite of disabilities, *languages* or *technological constraints*.
Practice of making websites	
Luján-Mora and Masri (2012)	With websites, the term traditionally refers to the *development* of websites accessible to all users who may want to access them, independent of the abilities or disabilities of the users. When websites are correctly designed and developed, all users can have equal access to information and functionality.
Chevalier et al. (2013)	Web accessibility is the inclusive *practice* of making websites usable by people of all abilities and disabilities.
Craven and Nietzio (2007)	… Web accessibility can also refer to the *design* of the web interface which, *according to recommended standards and guidelines*, should be presented in a way that can be interpreted by as wide a group of user as possible and by any kind of assistive technology.

As it can be seen from the analysis, a variety of different definitions of web accessibility can be found. Some definitions are more extensive than others, as they refer to more concepts than others. For example, the technology people are using. The positive thing is that there are not many discrepancies between the definitions, as almost all of them refer to the users, particularly disable users, being able to access and use the websites. Many of the definitions refer to the users' technology, with some explicitly referring to assistive technology. Many definitions refer to the characteristics of the websites as well as what users should be able to do. In addition, some definitions refer to the characteristics of the situations of use as well as the practice of creating accessible websites. Although, there is not much of a conflict between the concepts of web accessibility from the different definitions, a complete definition that takes together all these concepts into a single definition is required in order to have a full picture of the components that are part of the concept of web accessibility.

From the analysis, a unified definition of web accessibility was formed based on the different concepts. The proposed definition reflects the strength of importance of the concepts, as expressed by the frequency of the concepts in the 157 definition. The importance of the concepts is depicted in the layers of an onion diagram (see Figure 1). The most important concepts are at the core of the onion and concepts of less importance at the skin of the onion.

From the onion diagram, the following definition of web accessibility is extracted: "*all people, particularly disabled and older people, can use websites in a range of contexts of use, including mainstream and assistive technologies; to achieve this, websites need to be designed and developed to support usability across these contexts*".

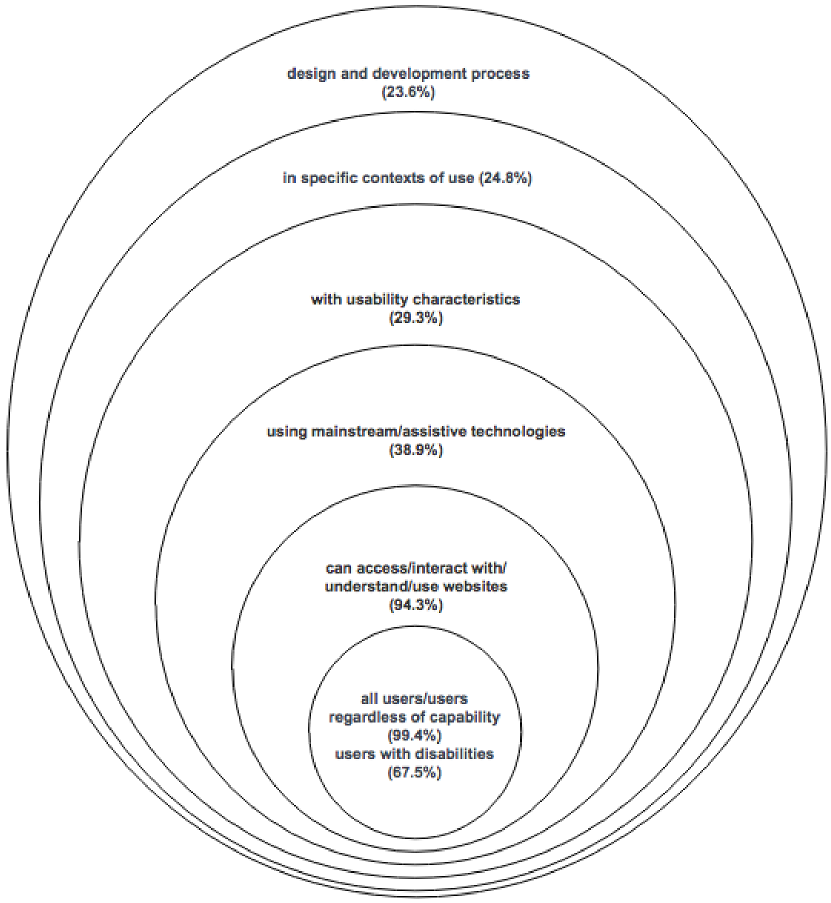

Figure 1. Six core concepts of web accessibility emerging from the analysis of 157 definitions.

The definition includes all the core components identified from the analysis. As the definition encompasses all the core components of the concept, the aim was to be coherent and comprehensive. Thus, for the component of what people can do, only the most frequent characteristic people do after accessing the web was selected.

Figure 2 shows the core components of the definition and how they related to each other. The figure can be used as guidance to define what researchers are controlling in studies that operationalise the concept, based on the unified definition of web accessibility.

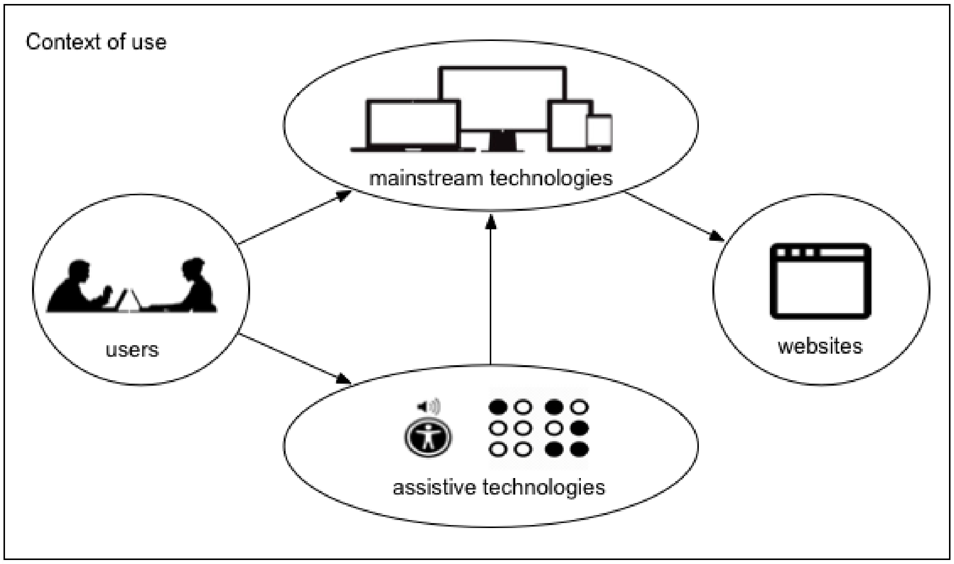

Figure 2. The core components of web accessibility.

3.4. Discussion

The aim of this study was to provide a better understanding what researchers consider to be the key components of the definition of web accessibility and propose a unified definition of web accessibility, drawing on all these components.

An analysis of 157 definitions found in the literature was performed, drawn from books, papers, standards, guidelines and online sources from an international selection of authors. The time period of the definitions was between 1996 and 2014. The analysis led to a unified definition of web accessibility. The propose definition encompasses all the core components of the concept. The onion diagram (see Figure 1) shows how the key components embedded within the definition fit together and make the definition clearer.

This study demonstrated that while there is often conflict and debate around the definition of web accessibility, most viewpoints and definitions can be reconciled into a single shared definition.

The proposed definition fits with the definitions proposed by the standards, as all of them refer to the use of websites by disabled users. The unified definition also fits with the definitions from standards about usability or qualities of usability (BSI, 2010; ISO 9241-171, 2008; Section 508, 1996), and the specific situations of use (BSI, 2010). The only concept that the unified definition does not comprehensively refer to in comparison to definitions from standards (BSI, 2010; WAI, 2005) is what people can do. However, this does not differentiate the viewpoints of the definitions from standards and the proposed unified definition. Moreover, the unified definition also refers to the technology that people use, whereas this concept is not mentioned by any of the definitions from standards.

The core components of the unified definition are in line with the viewpoints of some of the most commonly mentioned definitions of web accessibility in the literature, all users (Letourneau, 1998; Mankoff et al., 2005; Sierkowski, 2002; Thatcher et al., 2002), disabled users (Mankoff et al., 2005; Petrie & Kheir, 2007; Slatin & Rush, 2002; Thatcher et al., 2002), qualities of usability (Petrie & Kheir, 2007; Slatin & Rush, 2002), situations of use (Mankoff et al., 2005; Petrie & Kheir, 2007; Sierkowski, 2002), technology (Letourneau, 1998; Sierkowski, 2002). The only viewpoint that is not comprehensively referred to by the unified definition is what people can do (Letourneau, 1998; Sierkowski, 2002; Thatcher et al., 2002).

Although there are a few accessibility definitions that do not directly fit with the proposed unified definition. A few definitions where accessibility comes first to allow users to access the content were proposed, whereas usability improves the interaction with the website (Akhter et al., 2009; Buzzi et al., 2009a; Buzzi, Buzzi, Leporini, & Akhter, 2009b; Mori et al., 2011). That can imply that accessibility is a precursor to usability, which is in line with "universal usability" (Shneiderman, 2000, 2002). This viewpoint suggests that all usability problems fall within the scope of accessibility problems, even though there are studies that demonstrated that actually, this is not the case (Petrie & Kheir, 2007; Rømen & Svanæs, 2008, 2012).

As it can be seen, the unified definition fits with most viewpoints of standards and commonly referred definitions of web accessibility. However, it is interesting to look at some of the studies conducted on the web with disabled users and how they relate to the unified definition.

A positive sign is that some of the largest studies conducted with disabled users (Coyne & Nielsen, 2001; Disability Rights Commission, 2004; André Pimenta Freire, 2012) use components of the unified definition. Coyne and Nielsen (2001) varied the user group, which included blind, partially sighted and a sighted controlled group. The measures were usability qualities (effectiveness, efficiency and satisfaction). The Disability Rights Commission (2004) varied the user group, which included blind, partially sighted, dyslexic, physical impaired and hearing impaired, and a comparison between blind and sighted users. In the latter comparison, the study also varied the website, based on their accessibility rating. The measures of the study were qualities of usability (e.g. effectiveness, efficiency, perceived task difficulty). Petrie and Kheir (2007) varied the user group (blind and sighted) and the website. Qualities of usability were used as a measurement, such as effectiveness (task success rate), the number of problems encountered, the severity of the problems found. André Pimenta Freire (2012) varied the user group (blind, partially sighted and dyslexic). The measures included qualities of usability, for example, effectiveness, perceived task difficulty or the number of problems encountered.

Looking into the proposed unified definition of web accessibility and how it fits with studies of the field, it can be observed that it is related with pragmatic qualities of users' experience on the web, such as effectiveness, efficiency and satisfaction. It is not disagreed that accessibility can affect users' hedonic experiences, for example, users' emotions, but it is mainly studied through the pragmatic qualities of users' experience. Thus, accessibility can be considered a facet of user experience.

The proposed definition can be applied directly to existing studies in the literature. It provides clarity and clarifies what researchers are controlling in their study and what qualities of users' experience are being measured. The unified definition can be the basis for grounding further research that explores the relationship between different components of web accessibility. The unified definition components will be used to relate each of the variables of the next studies of this thesis.

Chapter 4. Empirical evaluation of the concurrent and retrospective verbal protocol for blind and sighted users

4.1. Introduction and Research Questions

This chapter presents the second study of this thesis. The study investigates which verbal protocol can be considered a better option in user-based studies with either blind or sighted users.

A review of the literature shows that CVP is the most preferred option for user-based studies with blind users on the web. However, there are some potential impacts and consequences of this choice, and may not reveal all the problems that blind users encounter on the web. RVP can be considered an alternative method (Chandrashekar et al., 2006; Strain et al., 2007), as it does not require users to think aloud as they perform the task. However, no research has been conducted comparing which verbal protocol can be considered a better option for blind users. With sighted users, there are some studies. However, the results of the studies do not bring light into which verbal protocol can be considered a better option, but also there are subject to methodological limitations.

The unified definition of web accessibility proposed in Chapter 3 was used to devise the study. Based on the core components of the concept (see Figure 2 in Chapter 3), the study manipulates the user group (blind or sighted). In addition, the study manipulates the protocol used by users to elicit problems. The measures included the problems users' encounter on the web and the effect of protocols on users. As the technology people are using is one of the key components of the unified definition of web accessibility, the problems in relation to users' technology (e.g. browser or assistive technology) were also considered as part of the problems users had.

This section presents the research questions of the study. The methodology, results and conclusion of the study are presented in the next sections.

The following research question was investigated:

- Which verbal protocol can be considered a better option for eliciting problems on the web for blind and sighted users?

In order to answer the research question, the following sub-questions were proposed, which can be grouped into three areas:

Effectiveness

- Does one protocol identify more problems than the other?
- Do blind and sighted participants identify the same number of problems with each protocol?
- Does one protocol identify more problems of a specific category than the other?
- Does one protocol identify more severe problems than the other?
- Do the two protocols identify the same problems?

Efficiency

- Does one protocol identify problems more rapidly?

Effect

- Does one protocol demand greater workload for participants, either blind or sighted?
- Does one protocol make participants more self-conscious than the other?
- Do participants prefer one method in comparison to the other?

4.2. Method

The study was a task-based user evaluation with blind and sighted participants using two different verbal protocols, CVP and RVP. A mixed design was used with the user group as the between-participant independent variable with two levels (blind or sighted) and with the verbal protocol as the within-participant independent variable with two levels (CVP or RVP). The dependent variables were the number of problems participants encountered in each protocol, the problems' severity ratings, the categories of problems revealed by each protocol, the number of problems identified per hour of evaluation time and participants' experience performing the protocol.

4.2.1. Participants

Sixteen participants took part in the study, eight blind screen reader users and eight sighted users. Six of the blind participants were men and two women. Ages ranged from 23 to 64 (*M* = 42.9, *SD* = 16.1). Three of the participants were congenitally blind while the remaining five lost their sight between the ages of 14 and 34. Sighted participants were selected to achieve as close matched sample as possible with the blind participants on gender, age, operating system used, web experience and web expertise. Thus, six of the sighted participants were men and two were women. Ages ranged from 22 to 55 years (*M* = 38.5, *SD* = 12.43).

Participants rated their experience and expertise on the web using a five-point Likert items (1 = "Very Low" to 5 = "Very Good"). The average rating of web experience for blind participants was 4.0 (*SD* = 0.9), whereas for sighted participants was 4.5 (*SD* = 0.5). Participants also rated their web expertise the same way. Blind participants' average rating was 3.8 (*SD* = 0.9), whereas for sighted participants was 3.6 (*SD* = 0.9).

All blind participants used screen readers to access computers and the web for home and work. Five participants used JAWS (running on Windows OS) and three used VoiceOver (running on Mac OS). The JAWS version varied from JAWS 12.0 to JAWS 15.0 (the latter being the latest version of JAWS when the study was conducted). Participants who used VoiceOver used the latest version of Mac OS Mavericks (the latest version of Mac OS when the study was conducted). Blind participants were asked to rate their experience and expertise of using screen readers on a five-point Likert item (1 = "Very Low" to 5 = "Very Good"). The average rating for experience and expertise using screen readers was 4.0 (*SD* = 0.5) and 3.9 (*SD* = 0.6), respectively.

Six participants used Mac OS (three blind and three sighted) and ten participants used Windows (five blind and five sighted). The majority of the blind participants who used Windows reported Internet Explorer as their primary browser and all of the participants who used Mac OS reported using Safari as their primary browser. Of the sighted participants, the ones who used Windows reported Google Chrome as their primary browser and one of them reported Internet Explorer. Of the ones using Mac OS, one user reported Google Chrome as their primary browser, whereas the other two reported Safari.

4.2.2. Equipment and Material

For participants who use the Windows OS, the study was conducted using a desktop computer running Window 8 with speakers, keyboard and 2-button mouse with scroll wheel. For participants, who use the Mac OS, the study was conducted using a MacBook Pro laptop running the

Mavericks Operating System connected with external speakers and a 2-button mouse with scroll wheel. In addition, blind participants were able to choose the screen reader software they were most familiar with, for example, JAWS, NVDA on Windows, or used VoiceOver version that comes with Mavericks OS on Mac. The screen reader software participants used was already declared during the recruitment process and all installations of the software were arranged properly before the arrival of the participants, in order to match their home or work environments.

It was preferred for participants not use their own equipment as it was easier to ensure that the equipment was in running order before the arrival of the participant. In addition, some of the participants may only have a desktop computer at home or at work. Thus, asking them to bring their own equipment would not be an option for them. Also, the sessions were recorded using screen recording software that was preinstalled on the computers used in the study, Morae 3.1 on Windows and ScreenFlow 4.0.3 on Mac OS. These recordings included audio, for the analysis of the verbal protocols, and screen activity for understanding the users' actions.

After each session participants completed the NASA TLX a subjective workload assessment questionnaire (Hart & Staveland, 1988). The NASA TLX assesses the subjective workload of a task using six workload components: mental demand, physical demand, temporal demand, effort, performance and frustration. Furthermore, a weighting scheme is included to consider individual differences. Participants are presented with every possible pair combination and asked to indicate which of the two workload components contribute more to the workload they experienced. Then participants are asked to give a rating for each workload component. The values obtained from the weighting procedure are used to weight the rating of each workload component. The NASA TLX is standard workload questionnaire that has been widely used in studies of interface design and evaluation (Hart, 2006).

At the end of the CVP session participants completed a questionnaire about the method using 5-point Likert items:

- Q1 (protocol interrupt): To what extent did thinking aloud during the task interrupt the flow of the task?
- Q2 (rating interrupt): To what extent did having to rate the problems for severity during the task interrupt the flow of the task?
- Q3 (protocol concentration): To what extent did thinking aloud during the task affect your concentration during the task?
- Q4 (rating concentration): To what extent did having to rate the problems for severity during the task affect your concentration during the task?
- Q5 (protocol real life): To what extent do you feel that thinking aloud during the task changed the way you did the tasks in comparison on how you might do it in real life?
- Q6 (protocol tiring): How tiring was it to do the think aloud during the task?

Participants answered Q1 - Q5 using scale: 1 = “Not at all” to 5 = “Very much”, and Q6 using a scale: 1 = "Not at all tiring" to 5 = "Very tiring".

At the end of both verbal protocols participants were asked to complete the following question:

- To what extent did thinking aloud during the task/replay of the task made you self-conscious about what you were doing?

Finally at the end of the session, participants were asked to select which one of the two verbal protocol they preferred conducting and to explain why they chose that preference.

4.2.3. Websites and Tasks

Four websites from different domains were used, a government website (www.gov.uk), a real estate website (www.rightmove.co.uk), an e-commerce website (www.boots.com) and a news/tv channel website (www.channel4.com).

The tasks used were:

- Gov.uk: Find how much it is going to cost to arrange a meeting to apply for a National Insurance number from your mobile phone number.
- Rightmove: Find a house to rent with a minimum of two bedrooms and a rent of no more than £1200 per month, near to a secondary school (a postcode was provided).
- Boots: Find the cheapest, five-star rated car seat for two year old child who weights 24kg.
- Channel4: Find which movie will be on Film4 at 9pm on the day after tomorrow.

The websites and tasks that were used investigate different design aspects of the websites, as they covered both navigation and input entries. They covered design aspects such as information architecture, navigation, content, headings, links, images, forms and tables.

I considered using websites on desktop or laptop computers, rather than mobile devices, to be consistent in the presentation of the websites evaluated. Many websites that are designed to adapt to the size of screen they are presented on, automatically change to fit users' device (Wroblewski, 2011). That means a website may look different on different mobile devices. Thus, having participants using the websites on mobile devices was not preferred for the purposes of the present study as it could potentially influence the results, as participants may have had interacted with different versions of the website, depending on the screen size of their mobile device.

In preparation for the evaluation, the tasks were first undertaken using JAWS and NVDA on Windows and VoiceOver on Mac OSX, to check that it is possible for screen reader users to be able to complete the tasks.

4.2.4. Procedure

The study took place in the Interaction Laboratory at the Department of Computer Science of the University of York and at the National Council for the Blind of Ireland (NCBI) in Dublin. Participants were first briefed about the study and were asked to sign an informed consent form. In order to avoid any conflicts between the technology and participants' preferences, participants were asked which browser they would like to use. Blind participants were also asked which screen reader they preferred and which version. They were also given the option to adjust the computer display, sound and related software to their preferences in order to match to their usual setup.

A demonstration on how to perform the verbal protocol the participant was about to conduct was first performed. Participants tried the protocol out using a practice website, which was not analysed in the study.

When participants were comfortable doing the appropriate verbal protocol, they were asked to perform each task. Depending on which protocol participants were using, they performed CVP or RVP. The verbal protocol approach that was used was based on the Boren and Ramey (2000) approach. I considered using this approach as the use of acknowledgements tokens makes the session more natural. In addition, a review of the literature shows that the two approaches, Boren and Ramey (2000) and Ericsson and Simon (1980, 1993) are equally applicable (Bruun & Stage, 2015; Olmsted-Hawala et al., 2010).

During the CVP condition participants thought out loud as they performed the tasks. When participants were quite for around more than 20 seconds, they were prompted with “What are you thinking about?” to remind them to vocalise their thoughts. However, the time intervals were not strict to 20 seconds in the case of blind participants, as there were occasions when blind participants were silent for more than 20 seconds as they were listening to the text from the website using the screen reader. For example, a participant may was looking for a specific link in a list of links that had more than one hundred links. Thus, the use of reminding prompts in this specific occasions relied on the researcher’s discretion. Participants were asked each time they encountered a problem to describe the problem and rate it for its severity using a four-point scale. Problems were considered everything that the participants felt that was a problem, whether that was caused by the website, the browser or the screen reader. The rating scale is based on the Nielsen’s severity rating for usability (Nielsen, 1994). However, the description of the problem was adopted to a user-centred description, as follows:

- Cosmetic problem (1): This problem on the website is making it slightly difficult to complete my task

- Minor problem (2): This problem on the website is making it difficult to complete my task
- Major problem (3): This problem on the website is making it very difficult to complete my task
- Catastrophic problem (4): This problem on the website makes my task impossible to complete

During the RVP condition participants performed the task in silence, then they reviewed the task as the video (or for blind participants, the audio) of the task was played back, immediately after participants completed each task. Participants were also allowed, depending on their preference, to control the video/audio using the spacebar button of the keyboard to pause and resume the flow, if they wanted more time to think out loud. Similar prompting and problem severity rating procedures were used in the RVP condition as in the CVP.

This procedure was repeated for each website. After doing two websites with one protocol participants were asked to complete the NASA TLX and the questionnaire about the method they used. The procedure was then repeated for the second verbal protocol. The order of the tasks and the verbal protocols were counterbalanced within each user group, to minimize practice and fatigue effects.

After completing both protocols, participants were asked to choose which one of the two verbal protocol they preferred and to explain why, as well as to complete a demographic questionnaire. Finally, participants were debriefed about the study and I answered their questions.

4.2.5. Data Analysis

The video recordings of each participant were reviewed, in order to extract the user problems. To support the problem matching, the problems identified were recorded using a structured problem report inspired by the ones that were used by Lavery, Cockton, and Atkinson (1997) and McDonald, Zhao, and Edwards (2015). The report included the context in which the problem occurred (website, page and in which protocol), a unique code number, a problem category and problem type assigned by the evaluator, the problem description as verbalised by the user and the user's assigned perceived severity rating.

The problem instances were categorised using a usability problems classification scheme. The classification scheme used was the one proposed by Petrie and Power (2012), which emerged from a corpus of 907 distinct problems identified on six websites by 30 users and 14 experts. Although other classification schemes exist (Van den Haak et al., 2003, 2004, 2007, 2009), the Petrie and Power classification scheme was preferred since it is similar to the one used in the studies by Van den Haak, but it is more explicit. For example, in the classification scheme used by Van den Haak et al. it is not clear when a comprehension problem occurs whether is about a content element or an interactive element. Petrie and Power classification scheme involves four different distinct problem categories: physical presentation content, information architecture, interactivity. Each of these problem categories is broken down into more specific problem descriptions. For example, the content category comprised of "too much content", "content not clear enough", "content not detailed enough", "content inappropriate or not relevant", "terms not defined" and "duplicated or contradictory content". An additional category was added to deal with the problems of incompatibilities with the website and the technology the user was using, named technology problems. To distinguish the differences between the content, information architecture and interactivity problems, we considered interactivity problems those that break the interaction of the user with the website, information architecture those that are related to the categorisation and the structure of the information between and within the pages and content problems those that are associated with the information in the pages.

Table 8 shows examples of each problem category from blind and sighted participants.

Table 8. Examples of each problem category from blind and sighted participants.

Problem Category	**Blind participants**	**Sighted participants**
Content	There is nothing about schools in the description of the house (P8)	The product description is limited. There is nothing about weight (P16)
Information architecture	The structure of the movies is confusing. I cannot understand which of the two times is the correct one for the movie (P5)	The option to filter by schools is very deep in the site (P13)
Interactivity	The input of the maximum number of bedrooms does not have a label (P1)	The group weight options in the filtering are not very clear (P15)

The next phase of the analysis involved identifying distinct problems. The problem instances checked if there were distinct problems, that is a problem that may have been encountered by more than one participant or by the same participant on more than one occasion on the same website in the same context.

Inter-coder reliability on the identification of problems was calculated on 10% of the video sessions. An additional evaluator, not involved in the study, independently extracted the problems from the videos. The reliability was calculated using the any-two agreement (see Equation 1) by Hertzum and Jacobsen (2001).

$$\frac{Pi \cap Pj}{Pi \cup Pj}$$

Equation 1. Any-two agreement by Hertzum and Jacobsen (2001)

The any-two agreement measures to what extent different evaluators agree on what problems the website contains. It is based on the number of problems the two evaluators have in common divided by the total number of problems they identified. P refers to the number of problems identified and i and j refers to the two evaluators. The conservative approach we followed in terms of the definition of the problem resulted in 100% agreement on the identification of user problems.

Inter-coder reliability (Landis & Koch, 1977) on the problem matching was calculated. This was achieved following the approach used by Frøkjær and Hornbæk (2008). Another researcher not involved in the study tried to match a random set of problems. The researcher received 10% of the problems together with the list of the matched problems from which these problems were not included. Then, for each problem, the researcher either matched it together with a problem in the list of the matched problems or noted that problem was not similar to any of the problems on the list of the matched problems. The analysis revealed satisfactory levels of agreement (K = 0.829).

Inter-coder reliability was also performed on the categorisation of 10% of the problems. A random sample of problems was provided to another researcher, who was asked to categorise the problems. The analysis revealed satisfactory levels of agreement, with K = 0.883 for the problem categories and K = 0.836 for the problem types.

4.3. Results

This section presents the results regarding the number of problems identified from the two verbal protocols, by both blind and sighted participants. Then I look at the distinct problems detected in both conditions for both blind and sighted participants. I will then discuss about the perceived severity of problems by participants in each protocol. Finally, I will discuss about the effect of the protocol on participants.

A total of 260 instances of problems yielded 136 distinct problems were identified, across both verbal protocols and both user groups.

4.3.1. Analysis of problem differences between the two verbal protocols and user groups

Problem Instances

To investigate whether one protocol identified more problem instances than the other and whether blind or sighted participants identified more problem instances, a two-way mixed ANOVA was conducted on the number of problem instances identified in each protocol condition and by blind and by sighted participants. The analysis revealed a significant main effect for the verbal protocol, $F(1, 14) = 6.93$, $p = 0.020$, $\eta^2_{partial} = 0.331$. The mean number of problem instances identified using CVP was 6.56 ($SD = 2.39$), whereas in RVP it was 9.69 ($SD = 4.27$). There was no significant main effect for user group, $F(1,14) = 3.06$, $p = 0.102$, $\eta^2_{partial} = 0.179$. The mean number of problem instances identified by blind users was 9.19 ($SD = 3.02$), whereas by sighted users it was 7.06 ($SD = 1.64$). Finally, there was no interaction effect between protocol and user group, $F(1, 14) = 0.00$, $p = 1.000$, $\eta^2_{partial} = 0.000$.

Problem categories

For the analysis of the problem categories, only the problems encountered by both user groups were considered (content, information architecture, interactivity), as blind participants did not encounter any physical presentation problems and sighted participants did not encounter any technology problems.

A three-way ANOVA (verbal protocol x user group x problem category) did not reveal any significant main effect for user group, $F(1, 14) = 3.19$, $p = 0.096$, $\eta^2_{partial} = 0.185$. The mean number of problem instances per problem category per protocol for blind users was 2.73 ($SD = 0.91$), whereas for sighted users it was 2.08 ($SD = 0.47$).

The analysis revealed a significant main effect for verbal protocol, $F(1, 14) = 5.30$, $p = 0.037$, $\eta^2_{partial} = 0.275$. The mean number of problem instances in CVP per problem category was 1.98 ($SD = 0.67$) per participant, whereas in RVP it was 2.83 ($SD = 1.33$).

The analysis also revealed main effect for problem category, $F(1.46, 20.42) = 41.07$, $p < 0.001$, $\eta^2_{partial} = 0.746$, with Greenhouse-Geisser correction. Post-hoc comparisons with Bonferroni correction indicated that the mean number of interactivity problems ($M = 4.53$, $SD = 2.16$) per participant per protocol was significantly higher than the mean number of content problems ($M = 1.25$, $SD = 1.00$) and the mean number of information architecture problems ($M = 1.44$, $SD = 0.87$). The other comparison (content and information architecture) was not significantly different.

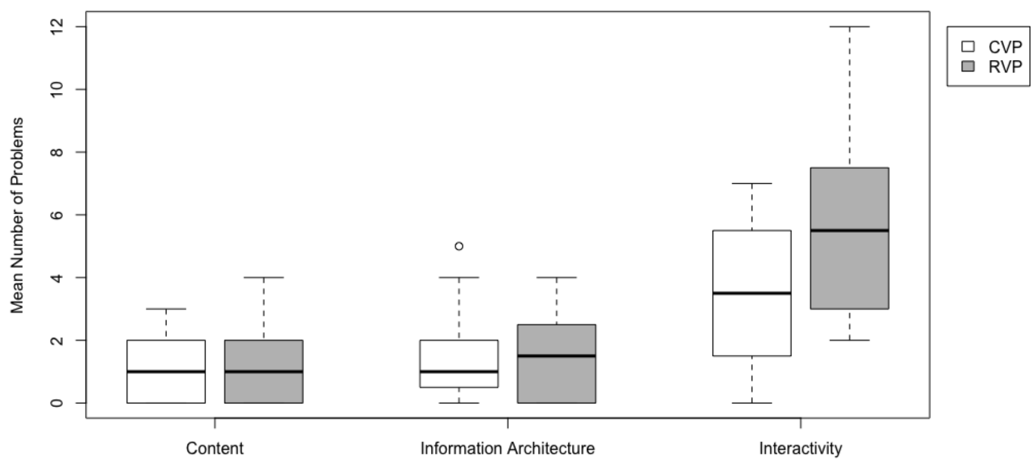

Figure 3. Boxplot showing the distribution of problem instances per problem category and verbal protocol.

There was a significant interaction between verbal protocol and the problem category, $F(2, 28) = 4.29$, $p = 0.024$, $\eta^2_{partial} = 0.235$. Figure 3 shows the distribution of problems for the three problem categories for both protocols. An analysis of simple effects showed that there was a significant difference between protocols for the interactivity problems, $F(1, 14) = 7.73$, $p < 0.05$, $\eta^2_{partial} = 0.36$. The mean number of interactivity problems identified using CVP was 3.38 ($SD = 2.36$), whereas in RVP it was 5.69 (*SD* = 3.05). None of the other comparisons was significantly different (see Table 9).

Table 9. Mean (SD) of participants' number of problems identified per problem category per protocol. *p < 0.05, **p < 0.01, ***p < 0.001.

Problem Category	**Verbal Protocol**		**F(1, 14)**	**p**	$\eta^2_{partial}$
	CVP	**RVP**			
Content	1.13 (1.09)	1.38 (1.26)	0.68	0.43	0.05
Information Architecture	1.44 (1.46)	1.44 (1.50)	0.00	1.00	0.00
Interactivity	3.38 (2.36)	5.69 (3.05)	7.73	0.02*	0.36

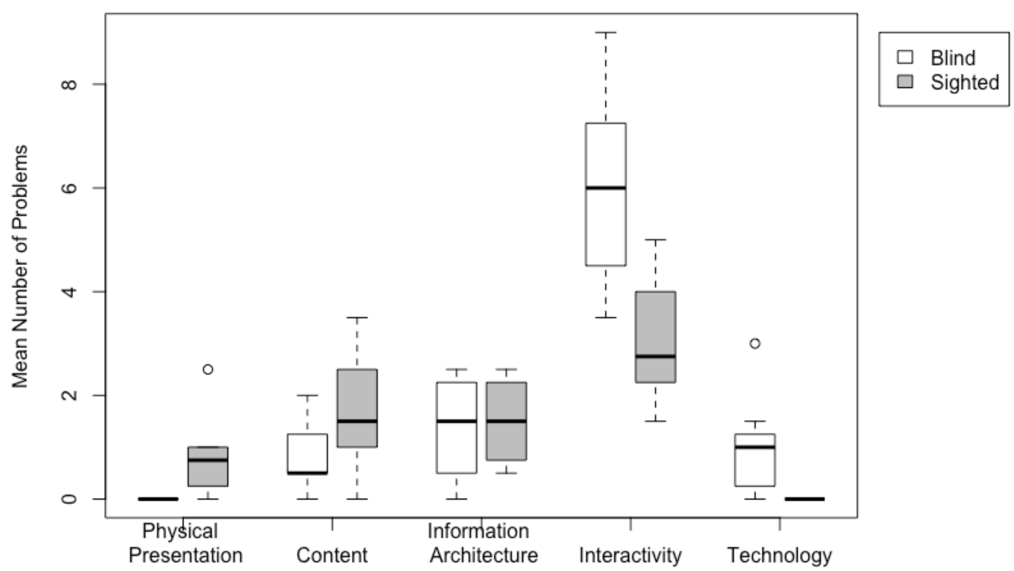

Figure 4. Boxplot showing the distribution of problem instances per problem category and user group.

There was also a significant interaction between user group and problem category, $F(2, 28) = 12.34$, $p < 0.001$, $\eta^2_{partial} = 0.468$. Figure 4 shows the distribution of problem instances per problem category by user group. An analysis of simple effects showed that there was a significant difference between blind and sighted participants on interactivity problems, $F(1, 14) = 13.53$, $p = 0.002$, $\eta^2_{partial} = 0.491$. The mean number of interactivity problems encountered by blind participants per protocol was 6.00 (*SD* = 1.91), whereas for sighted participants it was 3.06 (*SD* = 1.21). The other comparisons were not significantly different (see Table 10).

Table 10. Mean (SD) of participants' number of problems identified per problem category per user group. *p < 0.05, **p < 0.01, ***p < 0.001.

Problem Category	**User Group**		**F(1, 14)**	**p**	$\eta^2_{partial}$
	Blind	**Sighted**			
Content	0.81 (0.65)	1.69 (1.13)	3.59	0.079	0.204
Information Architecture	1.38 (0.95)	1.50 (0.85)	0.08	0.786	0.005
Interactivity	6.00 (1.91)	3.06 (1.21)	13.53	0.002**	0.491

Further examination of the interactivity problems showed that there were interactivity problems that encountered only by blind participants and not by sighted participants. These problems included labels missing on interactive elements (*N* = 5), lack of feedback on user actions (*N* = 3), links that lead to external sites without warning (*N* = 2), and input formats not clear (*N* = 6). In addition, there were interactivity problems that were encountered more frequently by blind participants than by sighted participants. These included instructions on interactive elements not clear (blind N = 40, sighted N = 16), options not complete (blind N = 9, sighted N = 5), interaction not as expected (blind N = 13, sighted N = 5) and elements not clearly identified as interactive or not (blind N = 6, sighted N = 2).

There was no interaction between user group and verbal protocol, $F(1, 14) = 0.03$, $p = 0.869$, $\eta^2_{partial} = 0.002$. Finally, there was no significant three-way interaction between problem category, verbal protocol and user group, $F(2, 28) = 1.13$, $p = 0.336$, $\eta^2_{partial} = 0.075$.

Problems overlap between the two verbal protocols for each user group

To investigate whether the two verbal protocols identified the same distinct problems and what percentage of problems was identified by each protocol, the distribution of distinct problems identified by each method and by both methods was calculated for blind and sighted participants separately. Figure 5 shows that for both user groups 27% of the distinct problems were found by both CVP and RVP, with slightly lower figure for sighted participants (23%) than for blind participants (31%). In total, RVP identified around 76% of the distinct problems, whereas CVP only identified 51% of the distinct problems.

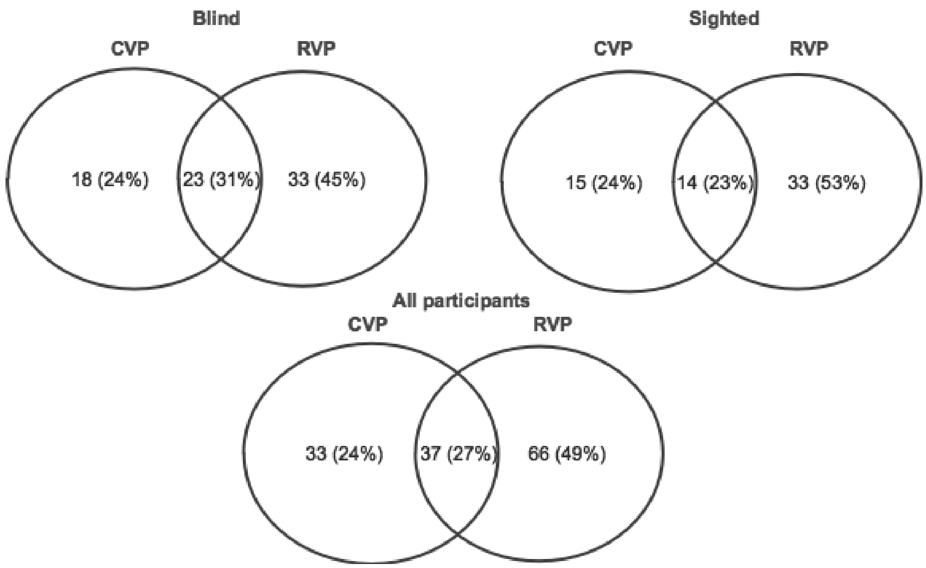

Figure 5. Numbers and percentages of distinct problems identified for each protocol for the two user groups and for all participants across the four websites.

Problems severity ratings between the two protocols and user groups

To investigate the severity of problem instances identified in the two protocols and by blind and by sighted participants, a two-way mixed ANOVA was conducted on the severity rating of the problem instances. There was no main effect for verbal protocol, $F(1, 14) = 0.62$, $p = 0.437$, $\eta^2_{partial} = 0.044$. The mean severity rating of problem instances for CVP was 2.30 ($SD = 0.55$), whereas for RVP it was 2.21 ($SD = 0.58$). There was no main effect for user group, $F(1, 14) = 0.00$, $p = 0.985$, $\eta^2_{partial} = 0.000$. The mean severity rating of problem instances for blind users was 2.25 ($SD = 0.53$), whereas for sighted users it was 2.26 ($SD = 0.55$). There was no interaction between protocol and user group, $F(1, 14) = 0.09$, $p = 0.640$, $\eta^2_{partial} = 0.016$.

To investigate whether the problems found by blind and sighted participants were rated more severely by one of the two protocols, the severity ratings of the problems that were found by both protocols were analysed. For blind participants, 23 problems were found by both protocols. The mean severity of these problems when found using CVP was 2.43 (SD = 0.98), whereas when found using RVP it was 2.12 (SD = 0.65). A paired sample t-test showed that there was no significance difference between these ratings from the two protocols, $t(22) = 1.81$, $p = 0.250$, $d = 0.246$. For sighted participants, 14 distinct problems were found by both protocols. The mean severity of these problems when found using CVP was 2.33 (SD = 0.93), whereas when found using RVP it was 2.40 (SD = 0.55). Again, a paired sample t-test showed that there was no significance difference between the ratings for the two protocols, $t(13) = -0.23$, $p = 0.814$, $d = -0.064$.

Efficiency for identifying problems between the two user groups

To investigate the efficiency of the two protocols an analysis of the number of problems identified per hour of evaluation time per participant was conducted. A two-way mixed ANOVA revealed that there was no main effect for protocol, $F(1, 14) = 1.62$, $p = 0.223$, $\eta^2_{partial} = 0.104$. The mean number of problems identified per hour of evaluation time per participant for CVP was 20.22 (SD = 12.40), whereas for RVP it was 17.84 (SD = 12.30). However, there was a main effect for the user group, $F(1, 14) = 30.17$, $p < 0.001$, $\eta^2_{partial} = 0.683$. The mean number of problems identified per hour for blind participants was 9.59 (SD = 4.36), whereas for sighted participants was 28.47 (SD = 9.96). Finally, there was no interaction between protocol and user group, $F(1, 14) = 0.66$, $p = 0.430$, $\eta^2_{partial} = 0.045$.

4.3.2. Protocol effect on participants

Protocols workload between the two user groups

To investigate the workload of undertaking the protocols for blind and sighted participants, an analysis of the overall NASA TLX was calculated. A two-way mixed ANOVA (protocol x user group) revealed a significant main effect for protocol, $F(1, 14) = 4.63$, $p = 0.049$, $\eta^2_{partial} = 0.249$. The overall mean NASA TLX score for CVP was 8.87 (*SD* = 3.64), whereas for RVP it was 11.11 (*SD* = 2.77). There was no main effect for user group, $F(1, 14) = 2.26$, $p = 0.155$, $\eta^2_{partial} = 0.139$. The mean NASA TLX score for blind users was 10.89 (*SD* = 1.80), whereas for sighted users it was 9.09 (*SD* = 2.88). There was also no interaction between verbal protocol and user group, $F(1, 14) = 0.58$, $p = 0.458$, $\eta^2_{partial} = 0.040$.

Participants' attitudes towards the two protocols

To investigate participants' attitude towards the two protocols, an analysis of the ratings on the six questions answered after completing CVP and the one question after completing either protocol was performed. Looking more specifically at the differences between the two user groups on the six questions (see Table 11), sighted participants found rating the severity of problems interrupted the flow of the task and their concentration more than blind participants.

Table 11. Mean (SD) of each question per user group and t-tests between user groups' attitude towards CVP. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

Question	**User Group**		**t (14)**	**p**	**d**
	Blind	Sighted			
Protocol interrupt (Q1)	2.13 (0.64)	2.25 (0.89	-0.32	0.751	-0.155
Rating interrupt (Q2)	1.50 (0.76)	3.00 (0.93)	-3.55	0.003**	-1.775
Protocol concentration (Q3)	2.00 (1.07)	2.25 (0.89)	-0.51	0.619	-0.255
Rating concentration (Q4)	2.00 (0.93)	3.00 (0.93)	-2.16	0.049*	-1.080
Protocol real life (Q5)	2.13 (1.64)	2.50 (1.31)	-0.51	0.621	-0.249
Protocol tiring (Q6)	1.75 (0.71)	1.75 (0.71)	0.00	1.000	0.000

One-sample t-tests were also conducted for each of the six questions for blind and sighted participants separately to investigate whether participants' ratings were significantly above the “not at all” point and significant different from the midpoint of the scale (“moderately”). The one-sample t-tests that were compared with the "not at all" value were one-tailed, whereas the other one-sample t-tests were two-tailed. Table 12 shows the results from the one-sample t-tests for blind and sighted participants.

Table 12. One-sample t-tests for blind and sighted participants' questions about CVP. *p < 0.05, **p < 0.01, ***p < 0.001.

User group	**Mean (SD)**	**Test value = 1**			**Test value = 3**		
		t (7)	**p**	**d**	**t (7)**	**p**	**d**
Protocol interrupt (Q1)							
Blind	2.13 (0.64)	4.97	0.001**	1.755	-3.86	0.006**	-1.365
Sighted	2.25 (0.89)	3.99	0.003**	1.410	-2.39	0.048*	-0.846
Rating interrupt (Q2)							
Blind	1.50 (0.76)	1.87	0.052	0.661	-5.61	0.001**	-1.983
Sighted	3.00 (0.93)	6.11	< 0.001***	2.160	0.00	1.00	0.000
Protocol concentration (Q3)							
Blind	2.00 (1.07)	2.65	0.017*	0.936	-2.65	0.033*	-0.937
Sighted	2.25 (0.89)	3.99	0.003**	1.410	-2.39	0.048*	-0.845
Rating concentration (Q4)							
Blind	2.00 (0.93)	3.06	0.009**	1.080	-3.06	0.018*	-1.082
Sighted	3.00 (0.93)	6.11	< 0.001***	2.160	0.00	1.000	0.000
Protocol real life (Q5)							
Blind	2.13 (1.64)	1.94	0.047*	0.685	-1.51	0.175	-0.533
Sighted	2.50 (1.31)	3.24	0.007**	1.146	-1.08	0.316	-0.382

Protocol tiring (Q6)							
Blind	1.75 (0.71)	3.00	0.010*	1.061	-5.00	0.002**	-1.768
Sighted	1.75 (0.71)	3.00	0.010*	1.061	-5.00	0.002**	-1.768

Blind participants found thinking out loud interrupted the flow of the task (Q1) and their concentration (Q3) significantly more than “not at all”, but significantly less than “moderately”. They found that rating problems for their severity interrupted their concentration significantly more than “not at all” but significantly less than “moderately” (Q4). Blind participants also found that performing the CVP was significantly different than the way they might do the tasks in real life (Q5). Further, they found that performing the CVP was significantly more tiring (Q6) than not performing it at all.

Sighted participants found that thinking aloud (Q1, Q3) and rating the problems for their severity (Q2, Q4) significantly interrupted the flow of the task and their concentration more than “not at all”. They also found that performing CVP changed the way they perform the tasks compared with real life (Q5) and that it was significantly more tiring (Q6) than not performing it at all. In comparison to the moderate midpoint, the results showed that sighted participants found that thinking aloud interrupted the flow of the task (Q1) and their concentration (Q3), although the interruption was significantly less than the midpoint of the scale. Also, they found performing CVP to be significantly less tiring (Q6) than the midpoint of the scale.

Participants were asked to rate how much thinking aloud during the tasks (for CVP) or during the replay of the task (during RVP) made them self-conscious about what they were doing (on a scale from 1 = "Not at all" to 5 = "Very much"). A two-way ANOVA revealed that there was no main effect for the protocol, $F(1, 14) = 0.13$, $p = 0.728$, $\eta^2_{partial} = 0.009$. The mean rating for CVP was 2.19 ($SD = 0.98$), whereas in RVP it was 2.06 ($SD = 1.18$). There was no main effect for user group, $F(1, 14) = 0.09$, $p = 0.768$, $\eta^2_{partial} = 0.006$. The mean rating for blind participants was 2.06 ($SD = 0.86$), whereas for sighted participants it was 2.19 ($SD = 0.80$). There was not interaction between protocol and user group, $F(1, 14) = 2.02$, $p = 0.177$, $\eta^2_{partial} = 0.126$.

One-sample t-tests were also conducted for the self-conscious question for blind and sighted participants for each protocol separately to investigate whether participants' ratings were significantly above the "not making them self-conscious at all" point and significantly different from the midpoint of the scale ("making them moderately self-conscious"). Table 13 shows the results from these one-sample t-tests. Blind participants found both protocols made them significantly more self-conscious about what they were doing than not doing them at all. However, when the results were compared with the midpoint value of 3, blind participants found that doing CVP made them significantly less self-conscious than the midpoint of the scale. Sighted participants found only that doing CVP made them significantly more self-conscious about what they were doing than not doing anything at all.

Table 13. One sample t-tests on ratings of self-consciousness of the two protocols, for blind and sighted participants. *p < 0.05, **p < 0.01, ***p < 0.001.

User group / protocol	**Mean (SD)**	**Test value = 1**			**Test value = 3**		
		t (7)	**p**	**d**	**t (7)**	**p**	**d**
Blind CVP	1.87 (0.83)	2.97	0.011*	1.049	-3.81	0.007**	-1.348
Blind RVP	2.25 (1.04)	3.42	0.006**	1.208	-2.05	0.080	-0.724
Sighted CVP	2.50 (1.07)	3.97	0.003**	1.403	-1.32	0.227	-0.468
Sighted RVP	1.88 (1.36)	1.83	0.056	0.645	-2.35	0.051	-0.829

Protocol preference between the user groups

Participants selected which of the two protocols they preferred undertaking. Five out of eight sighted participants preferred CVP and three preferred RVP, whereas of the eight blind participants four preferred CVP and four preferred RVP. There was no statistically significant association between user group and protocol preference as assessed by Fisher's exact test, $p = 1.000$.

4.4. Discussion

This study investigated the use of two verbal protocols for conducting evaluations in terms of effectiveness, efficiency and the effects they had on blind and sighted participants.

In terms of effectiveness, the results indicate that RVP is more effective than CVP. The results indicate that RVP produced significantly more problems overall. This result is in line with the findings of the studies conducted by Van den Haak et al. (2003, 2004), where it was also found that more problems were reported by users during the RVP condition. A possible explanation why RVP revealed more problems than CVP is that participants had more time to verbalize problems during RVP condition. Unlike CVP, participants in RVP verbalise their thoughts after finishing the task, which means that they can be fully focused on verbalising their thoughts. It seems that the double work, performing the task and thinking aloud the same time, may affect participant' ability to verbalise which prevents them from verbalising all the problems they encounter during the task.

Regarding the category of problems revealed, the results showed that interactivity category problems was the most frequently reported problem category, with RVP revealing more interactivity problems than CVP. A possible explanation for this may be that interaction with the page content requires more attention from users whilst simultaneously carrying out another action; thinking aloud and performing the task the same time. This may have had an effect on the reporting of interactivity category problems.

The study also showed that blind participants encounter more interactivity problems than sighted participants. The difference in frequency in interactivity problems between blind and sighted participants comes from several sources. There were interactivity problems that only encountered by blind participants, for instance, the lack of feedback on user actions and system progress, missing labels on interactive elements, and links that lead to external sites without warnings. There were also types of problems that were encountered by both user groups but which blind participants encountered more frequently than sighted participants. For example, instructions on interactive elements not clear and options not complete.

In addition, RVP identified more distinct problems than CVP for both blind and sighted participants. Comparing the two protocols in terms of whether they identify the same problems, it was found that only 27% of the distinct problems were identified by both protocols. The overlap of the problems between the two protocols in this study is similar with the overlap found in two of the studies by Van den Haak et al. (2007, 2009), between 34% and 38%. In the other two studies (Van den Haak et al., 2003, 2004) found a slighter higher overlap between the two protocols, between 47% and 56%. A limitation of the studies by Van den Haak et al. is that it did not specifically report the overall between CVP and RVP in three of the studies (Van den Haak et al., 2004, 2007, 2009), instead it provided a range of overlap between two out of the three protocols included in the studies. Moreover, the overlap included problems reported by both experts and users, whereas in this study only the problems reported by users were considered.

In this study, RVP revealed 76% of the total number of distinct problems, whereas CVP revealed only 51% of the total distinct problems, with a similar figure for both blind and sighted participants. The figure of problems revealed by RVP in this study is very similar with the figure reported in the study by Van den Haak et al. (2003), with RVP revealing 69% of the total distinct problems. Moreover, no support was found that participants in either user group perceived the problems found on both protocols more or less severe in one protocol in comparison to the other.

In terms of the efficiency of the two protocols for identifying problems, no support was found that the two protocols differ. However, it was found that there was a difference between the two user groups, with sighted participants identifying nearly three times the number of problems per hour of evaluation compared to blind participants. This result is not surprising as blind users interact with the website differently from sighted users and typically take longer to complete the tasks. Blind participants took approximately three times as long to complete tasks as the sighted participants, result very much in line with the results from the Digital Rights Commission investigation of web accessibility (Disability Rights Commission, 2004).

In terms of the effects of the protocols on participants, the NASA TLX showed that RVP demanded more workload overall than CVP for both blind and sighted participants. This result was unanticipated as it was expected that RVP would be less demanding than CVP, as participants would not have to think aloud and perform the task simultaneously. A possible explanation for this might be that because RVP takes at least twice as long compared to CVP, it may have had an effect on participants' workload, even though participants had to perform a single task, perform the task or verbal protocol, each time. There was a number of differences between blind and sighted participants on their perception of the two protocols. Sighted participants found rating the severity of the problems more disruptive than blind participants. However, comparing the ratings of the blind and sighted participants separately against "not at all" disruptive and "moderately" disruptive points revealed that both groups did find that CVP interrupted the flow of the task and concentration somewhat. Comments from blind participants on this disruption included:

"when I think aloud I may miss what JAWS is talking to me and I may forget what I was doing and where I was" (P2)

"when I was trying to find things, I had to think aloud and interrupted my concentration. it is difficult and sometimes frustrating" (P5)

"I was not listening 100% on JAWS…there is a lot of processing information I had to use a lot of senses" (P3)

These comments highlight how blind participants found thinking aloud interrupted their concentration and may cause them miss output from the screen reader. It was difficult for them to think aloud while they were trying to process the output of the screen reader and perform the task at the same time.

Comments from sighted participants on the disruption included:

"… trying to think aloud did interrupt the flow of the task" (P13)

"…by verbalizing my thoughts through process I assumed I was missing something (P15)"

"…my concentration was less focused than normal" (P12)

These comments highlight how sighted participants found that thinking aloud interrupted the flow of the task and their concentration.

Although participants found CVP somewhat interrupting the flow of their tasks, there was no difference between user groups in preference to the protocols. Participants were also asked to explain their choice. Comments from participants who preferred CVP included:

"because it's quicker" (P1, blind)

"it's in real time… beneficial at the time" (P3, blind)

"It was my normal way…I talk to the screen regularly" (P9, sighted)

Comments from participants who preferred RVP included:

"I found [RVP] more easy to follow during the replay of the task" (P2, blind)

“it was easier to do the tasks [in RVP] in silence you were able to concentrate more on what you were doing...RVP was easier because it was easier to listen to VoiceOver” (P4, blind)

“thinking aloud during the task was hard…forgetting what I was doing…it was a distraction…RVP was easier but demanded more time” (P8, blind)

“easier to do one thing at a time” (P12, sighted)

“I think it was easier to articulate your thoughts after having done them…Generally, it was just easier to do less multi-tasking” (P14, sighted)

The comments show that some participants found it easier to perform RVP, as it did not interrupt them, especially blind participants who had to process the output of the screen reader in addition to performing the protocol. However, other participants preferred CVP because it was quicker compared to RVP.

Based on the results, RVP is more profitable than CVP, particularly if the interest is on interactivity problems for eliciting problems on the web for both blind and sighted users. The result of this study guided the verbal protocol selected in the subsequent study of this thesis.

Even though the study provided a better understanding of the differences between the two verbal protocols in regard to which protocol is better to use for eliciting problems on the web it is subject to some limitations.

The conduction of the study in two different places is a potential threat to the internal validity of the results. The study was a split-site study, due to the difficulties of recruiting blind participants. As the study is targeting a particular user group, which is relatively small, recruiting participants that match the study specification was a challenging task.

Recruiting blind participants is a quite known issue in the research community and is one of the reasons of the small number of participants in research studies, with half of the studies having six or less blind participants (see Appendix A.). Due to the difficulties of recruiting blind users, half of the evaluations from blind users were conducted at the Interaction Laboratory at the Department of Computer Science of the University of York and the other half at the NCBI in Dublin, which helped me recruit participants. The first limitation of this split-study is the political differences that may influence participants elicitation of problems as one of the website was the public sector information website of the United Kingdom. The second limitation of this split-study is the differences in the environment that the study took place that may have had an impact on users' collected measures. To mitigate any influences the spit-study may have had, the evaluations in both places were moderated and it was tried to achieve as close as possible match environment. The studies were conducted in a room, with only the evaluator (author of the thesis) and the participant, using the same technologies, following the same study procedure wherever the study took place.

The external validity (i.e. the results generalised when users' use other systems) is also threatened in this study. The current study was focused on desktop websites rather than on mobile devices, in order to have a consistent presentation of the page content. On mobile devices, the websites adapt the page content to fit into their screen size. There is a possibility of a website to be presented differently on various mobile devices due to the different screen sizes, which may potentially influence the results, as the participants would not have a consistent presentation of the website content. To avoid any influences the different mobile devices may have had on the data measured, it was preferred to avoid using mobile devices so that the page content is presented consistently to users. It would have been interesting though to know whether the findings of the current study can also be generalised to when users use websites on mobile. Further research regarding the differences between the two verbal protocols for eliciting problems on the web on mobile websites would be worthwhile.

The low number of websites evaluated is another limitation of this study. Having more websites was not preferred as each website evaluation with blind users required around an hour to conduct. To mitigate any fatigue effects that users may have had the number of websites selected was low. The four websites selected varied in content and structure and involved both navigation and data input in covering different aspects and features of websites.

In the present study participants did not use their own computers, which may introduce a threat to ecological validity. Although it is not disagreed that it would have been better if participants used their own equipment to do the study, any impacts that the use of not their own equipment mitigate as participants were given the option to adjust the equipment to their usual setup. Also, participants first performed a practice task, that was not analysed in the study, to become comfortable doing the verbal protocol but also help them to become familiar with the equipment they were using. In addition, it was preferred for participants not to bring their own equipment as I was not going to be able to check that the equipment was in running order before the arrival of the participant. Also, some participants may have a desktop computer, which would have made their participation difficult. Also, screen recording software was preinstalled on the computers used in the study that capture the session. For these reasons, it was preferred for the participants not to use their own equipment.

Another limitation is in regard to the analysis of the dataset be over-tested. The dataset consisted of multiple measures from the same participant. For example, the number of problems users found, participants' workload, participants' preference. Although it seems reasonable to analyse each of these measures separately, there is a possibility that there may have been a relationship between the measures that actually push a test into a significance (Cairns, 2007). However, this was a risk that needed to be taken into consideration due to the pragmatic limitations of recruiting blind participants. When working with specific target groups, especially ones with a small population, it is common to collect data for more than a single measure (Coyne & Nielsen, 2001; Disability Rights Commission, 2004; André Pimenta Freire, 2012). Although, collecting data for a single measure would have made the analysis sound, due to the difficulties of recruiting participants it was preferred not to.

The own workload of NASA TLX questionnaire for blind users is another limitation that needs to be taken into consideration. The NASA TLX is comprised of two parts, the assessment of the six workload components and the weighting procedure. Before participants complete each step, participants read the instructions for each step (see Appendix E.). Completing the NASA TLX questionnaire can impose its own workload, particularly in the case of blind users as the questionnaire instructions were read to them by the evaluator (author of this thesis) compared to sighted users that read the instructions themselves. This approach was preferred for two reasons. First, previous research showed that a computer-based version of the NASA TLX could impose more workload itself than a paper-based version (Noyes & Bruneau, 2007). Also, having a computer version of the questionnaire (instructions and answers) could potentially lead to some problems when users try to complete the questionnaire via a screen reader, that may influence the measures collected for the study. Therefore, the paper-based approach was preferred. There is, however, a possible limitation of how the paper-based questionnaire was used in the study. The paper-based approach may have imposed higher workload on blind users compared to sighted users, as the instructions were read to them compared to sighted users that read the instructions themselves. Looking back to the study design it would have been better if the questionnaire was read to both user groups so that they had the same experience. This limitation itself provides us with further insights into areas that researchers should pay more attention when designing studies with both blind and sighted users.

4.5. Conclusions

This study has compared two verbal protocols, CVP and RVP, with blind and sighted participants. The two protocols were compared in terms of effectiveness and efficiency for identifying user problems on websites and the effect they have on participants. The study provides insight in terms of which verbal protocol is appropriate for use in studies with both blind and sighted participants.

The study has shown that RVP outperforms CVP in terms of effectiveness but is not more efficient than CVP for identifying user problems on the web. RVP identifies more distinct problems and problem instances than CVP for both blind and sighted participants.

Further, the study demonstrated that there was quite a low overlap in the problems between the two protocols identified for both blind and sighted participants. This result is consistent with previous studies (Van den Haak et al., 2007, 2009). In addition, RVP identified three-quarters of the total number of distinct problems, whereas CVP only identified half of the distinct problems. It was also shown that RVP revealed more interactive category problems than CVP.

Even though RVP created a significantly higher workload for participants and CVP was perceived as being somewhat disruptive of the flow of the task, there was no clear preference amongst participants for one protocol over the other, so these did not strongly differentiate between the protocols.

The study has provided a better understanding of the differences between the two verbal protocols. The results of the study can guide which verbal protocol can be considered a better option. RVP can be considered a better option in user-based studies, particularly if the interest is in interactivity problems. However, for studies interested in content or information architecture problems, either protocol is appropriate. For example, early prototypes of websites that do not require any interaction with the page content. It is believed that is the first study to compare the two verbal protocols with blind participants and it has provided insights into the differences between the two verbal protocols.

The result of this study guided the verbal protocol method that was used in the next study of this thesis, as the focus of the subsequent study was on identifying the problem differences between blind and sighted users on the web considering all the spectrum of problems users may encounter, including interactivity problems. Thus, for the purposes of the next study RVP was considered a better option.

Chapter 5. Empirical study of the problems between blind and sighted users on the web

5.1. Introduction

Previous research has indicated that there known overlaps and differences between the problems blind and sighted users have on the web (Petrie & Kheir, 2007). However, there is little information of what these problems are as well as what causes them. The aim of the present study is to investigate the problem differences and similarities between blind and sighted users on the web.

The unified definition of web accessibility proposed in Chapter 3 was used as the theoretical construct of the concept to be operationalised. Based on the key components of the concept (see Figure 2 in Chapter 3), the study manipulates the user group (blind or sighted users). The measures included qualities of usability, such as the problems users encounter on the websites and users' effectiveness. Problems in relation to the users' technology (e.g. browser, screen reader) and the websites' content were also collected, as users' technology is one of the key components of the unified definition.

Three online websites with related tasks to the websites' content were used in the experimental setting. Based on the results in Chapter 4, RVP was selected as the most appropriate verbal protocol to use, as the focus of the present study was to explore the problems users encounter on the web.

The following research question was investigated:

- What are the problem similarities and differences between blind and sighted users on the web?

In order to answer the research question, the following sub-questions were proposed:

- Does one user group encounter more problems than the other?
- Does one user group encounter more problems of a specific category/type than the other?
- Do the two user groups encounter the same problems?
- Does one user group perform the tasks more successfully than the other?

5.2. Method

5.2.1. Study Design

The study used a mixed-factor design, with user group as the one between-participant independent variable with two levels (blind or sighted) and the websites as the within-participant variable with three levels (Reed, theAA, ToysRUs). The dependent variables were qualities of usability. For example, the number of problem instances participants encountered, the number of problem instances for each problem category, the problems' severity rating and participants' task success rate.

5.2.2. Participants

A total of 24 participants took part in the study, 12 blind screen reader users and 12 sighted users. Nine of the blind participants were men and three were women. Ages ranged from 24 to 64 (M = 44.7, SD = 14.8). Five of the participants were congenitally blind while the remaining seven lost their sight between the ages of 14 and 42. Due to the limited participant pool, 4 out of the 12 blind participants, were participants that took part in the previous study of this thesis (see Chapter 4).

Sighted participants were selected to achieve as close as a matched sample as possible with blind participants on gender, age, operating system used, web experience and web expertise. Thus, nine of the sighted participants were men and three were women. Ages ranged from 26 to 70 (M = 46.7, SD = 15.6).

Participants rated their experience and expertise on the web using a five-point Likert item (1 = “Very low” to 5 = “Very good”). The average rating for web experience for blind participants was 4.5 (SD = 0.5), whereas for sighted participants was 4.0 (SD = 0.9). Participants also rated their web expertise the same way. Blind participants’ average rating was 4.1 (SD = 0.7), whereas for sighted participants was 3.5 (SD = 0.9).

All blind participants used screen reader software to access a computer on the web for home and work. Eight participants used JAWS, one NVDA (both running on Windows OS), and three used VoiceOver (running on Mac OS). The JAWS version used varied from JAWS 14.0 to JAWS 16.0 (the latter being the latest version of JAWS when the study was conducted). The participant who used NVDA used the latest version of this software (2015.1) when the study was conducted. Participants who used VoiceOver used the latest version of Mac OS Yosemite (the latest version of Mac OS when the study was conducted). Participants were also asked to rate their experience and expertise using screen readers on a five-point Likert item (1 = “Very low” to 5 = “Very good”). The average rating for experience and expertise using screen reader was 4.4 (SD = 0.5) and 4.1 (SD = 0.7), respectively.

Eighteen participants used Windows (nine blind and nine sighted), whereas the remaining six participants used Mac OS (three blind and three sighted). From the blind participants who used Windows, five mentioned Internet Explorer and four Firefox as their primary browser, and all the participants who used Mac OS reported Safari as their primary browser. From the sighted participants who used Windows, six mentioned Chrome and three Firefox as their primary browser. From the sighted Mac OS users, two mentioned Safari and one Chrome as their primary browser.

5.2.3. Equipment and Material

For participants who used the Windows OS, the study was conducted using a desktop computer running Windows 8 with speakers, keyboard and a 2-button mouse with scroll wheel. For participants who used the Mac OS, the study was conducted using a MacBook Pro running the Yosemite OS, with speakers and a 2-button mouse with scroll wheel. In addition, blind participants were able to choose the screen reader software they were most familiar with, for example, JAWS, NVDA or used the VoiceOver version that comes with Yosemite on Mac. The screen reader software that participants used, was already declared during the recruitment process and installation of the software was already arranged properly before the arrival of the participants, in order to match their home or work environment.

It was preferred participants not to use their own equipment as I wanted to ensure that the equipment is in running order before the arrival of the participants. In addition, some of the participants may not have a laptop at home or at work, hence, asking them to bring their own equipment was not going to be an option for them. Also, the sessions were recorded using screen recording software, Morae 3.1 on Windows and ScreenFlow 4.0.3 on Mac OS, that was preinstalled on the computers used in the study. These recordings included audio for the analysis of the verbal protocol and screen activity for understanding the users' actions.

5.2.4. Websites and Tasks

Three websites from three different domains were used, a job recruitment, an automotive and an e-commerce website.

The tasks used were:

- *Reed*: How many jobs are listed for graduates in engineering, which are full-time with salary more than £20k a year and were posted in the last week?
- *theAA*: Find a used car to buy. You want one, which has automatic transmission, it's less than 7 years old, fuel type is hybrid and price is between £1000 and £20,000. What is the cheapest car that meets these requirements?
- *ToysRUs*: What is the cheapest scooter, which is suitable for children aged between 5 to 7, and has been reviewed by other costumers?

The inclusion criteria for websites and tasks was to cover different design aspects, such as navigation between and within pages, headings, links, images, forms and content. Moreover, the results of the previous study (see Chapter 4) showed that both user groups encounter many interactivity problems on the web. For this reason, websites and tasks that require interaction with the page content were selected. Shopping websites or websites that require users to interact with filtering options and navigate through the page content were considered as an appropriate type of websites, as they cover many different interactivity features. For example, interact with links and form elements, such as buttons, checkboxes, input elements. Also, users would need to navigate through the page content to find particular information. The inclusion of websites that contain multimedia content (e.g. audio or video) was not chosen, as the problem differences related around this specific aspect between the two user groups are already known and solutions are already suggested by the accessibility guidelines (Caldwell et al., 2008). For example, in the case of video content on a page, audio description should be provided.

All sessions were conducted on desktop computers. The use of mobile devices was not considered in the present study. Websites that are developed for mobile devices adapt the page content to fit into the mobile screen size. However, the presentation of the page content is often depended on the screen size and that can vary depending on the mobile. To have a consistent presentation of the website all evaluations were conducted on a desktop or laptop device.

In preparation of the evaluation, the tasks were first undertaken using JAWS and NVDA on Windows and VoiceOver on Mac OS, to check that it is possible for screen reader users to be able to complete the tasks.

5.2.5. Procedure

The study took place in the Interaction Laboratory at the Department of Computer Science of the University of York. Participants were briefed about the study and were asked to sign an informed consent form. To avoid any conflicts between the technology and participants' preferences, participants were asked which browser they would like to use. Blind participants were also asked which screen reader they preferred and which version. They were also given the option to adjust the computer display, sound and related software to their preference in order to match their usual setup.

A demonstration on how to perform the verbal protocol (RVP) was performed by performing a task on a demonstration website. For blind participants, the task was performed using the screen reader they selected to use. Participants then tried the protocol using a practice website that was not analysed in the study.

When participants were comfortable doing RVP, they were asked to perform each task. The verbal protocol approach followed was the Boren and Ramey (2000) approach. I considered using this approach as is makes the interaction between the participant and the evaluator more natural. During the replay of the video, if the participants were quiet for more than 20 seconds, they were prompted with “What are you thinking about?” to remind them to vocalise their thoughts. Each time participants encountered a problem, they were asked to describe the problem.

The order of the tasks was counterbalanced within each user group, to minimize practice and fatigue effects. After completing all the tasks, participants were asked to complete a demographic questionnaire. Then participants were debriefed about the study and their questions were answered. Finally, any information that was necessary for the compensation of participants’ time was collected.

5.2.6. Data Analysis

The total amount of video recordings of the evaluations sessions was 23 hours and 24 minutes. The video recordings of each participant were reviewed, in order to extract the user problems. The same approach as described in section 4.2.5 was performed for the problem matching and categorization of the problems found.

Inter-coder reliability on the identification of problems was calculated on 10% of the video sessions. An additional evaluator, not involved in the study, independently extracted the problems from the videos. The reliability was calculated using the any-two agreement by Hertzum and Jacobsen (2001). The conservative approach followed in terms of the definition of problems resulted in a 100% agreement on the identification of user problems.

Inter-coder reliability (Landis & Koch, 1977) showed satisfactory levels of agreement for the problem matching (K = 0.867) and for the categorisation of problem categories and problem types, K = 0.880 and K = 0.834, respectively.

5.3. Results

This section presents an analysis of the data collected from blind and sighted participants regarding the problems they encountered on the three different websites tested. The section begins with a presentation of participants' task success. Then it presents an overview of the differences in the number and problem categories in order to identify specific differences and areas of interest for further analysis. Finally, for each problem category encountered by users, the differences and similarities between the two user groups are examined in detail, to provide an understanding of how the participant groups differ.

5.3.1. Participants' task success

In total 72 tasks were attempted across the three websites by the 24 participants. A one-way ANOVA on participants' task success rates showed a significant difference between the two user groups, $F(1, 22) = 5.25$, $p = 0.032$, $\eta^2_{partial} = 0.193$, with sighted participants being more successful in completing the tasks, compared to blind participants. Table 14 shows the task success rates for blind and sighted participants, with the percentage of task succeeded and failed for each user group.

Table 14. Task success rates for blind and sighted participants.

User group	Task succeeded	Task failed
Blind	53%	47%
Sighted	75%	25%

5.3.2. Frequency of problems encountered

The analysis revealed 526 problem instances encountered by both user groups. Blind participants encountered 381 problem instances, whereas sighted participants encountered 145 problem instances across the three websites.

In order to compare the problems between the user groups, it is important to identify where there are key differences between the user groups.

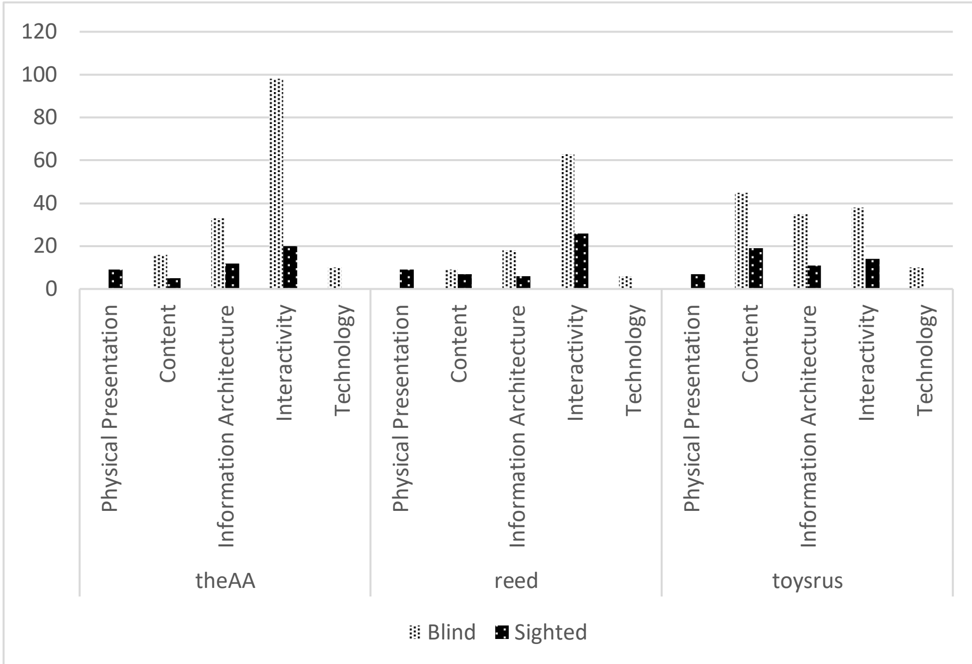

Figure 6. Frequency of problem instances for each problem category per website by user group.

Table 15. Frequency of problem instances for each problem category per website by user group.

Website and user group	**Physical Presentation**	**Content**	**Information Architecture**	**Interactivity**	**Technology**
theAA					
Blind	-	16	33	98	10
Sighted	9	5	12	20	-
Reed					
Blind	-	9	18	63	6
Sighted	9	7	6	26	-
ToysRUs					
Blind	-	45	35	38	10
Sighted	7	19	11	14	-

Specifically, we wanted to know if there were differences in the number of problems encountered by each user group and if there were differences between the number of problems found in each category or between the websites. Figure 6 and Table 15 present the number of problems distributed across the websites and the five main problem categories.

As is presented in Table 15 there appear to be some differences between the problem categories encountered on the web between the two user groups. There are problem categories distinct to each user group. For example, physical presentation problems category was distinct to sighted users, whereas technology problems category was distinct to blind users.

A three-way ANOVA (problem category x website x user group) was performed to assess whether there was a statistical difference between the two user groups regarding the problem category instances encountered on the three websites. The data violated the assumption of homogeneity of variance. To equate group variances, a square root transformation was performed (Howell, 2012). The homogeneity of variance of the transformed data was determined using the Levene's test. Untransformed values are displayed in the figures to aid interpretation.

The analysis revealed a significant main effect for the problem category, $F(2, 44) = 27.02$, $p < 0.001$, $\eta^2_{partial} = 0.551$. Means and standard deviations for each problem category are presented below.

Table 16. Mean (SD) of participants' problems in each problem category for untransformed and transformed data.

Problem Category	Untransformed	Transformed
Content	1.40 (1.22)	0.86 (0.56)
Information Architecture	1.60 (1.51)	0.98 (0.60)
Interactivity	3.60 (2.88)	1.61 (0.78)

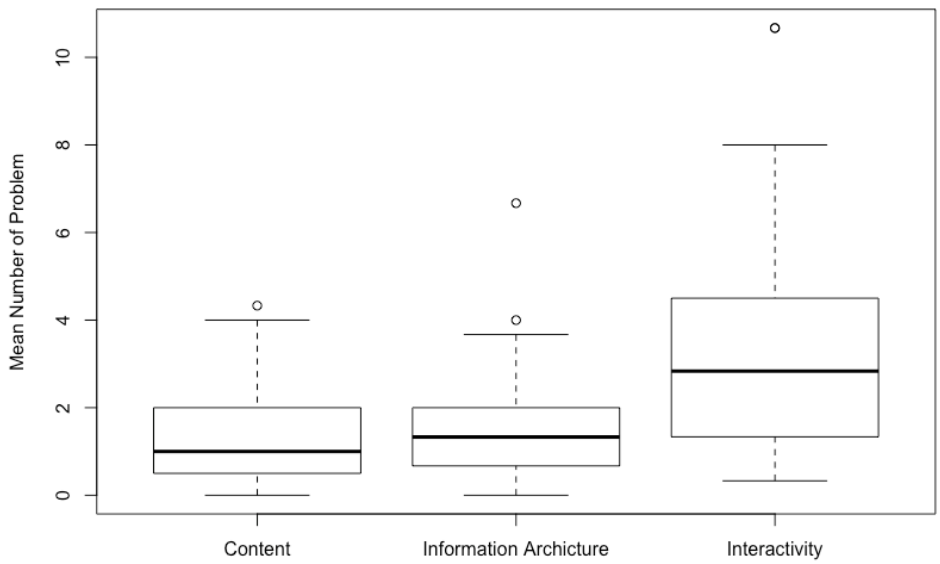

Figure 7. Boxplot showing the distribution of problem instances per problem category (untransformed data).

Post-hoc comparisons with Bonferroni correction indicated that the mean number of interactivity problem instances (untransformed data: M = 3.60, SD = 2.88, transformed data: M = 1.61, SD = 0.78) was higher than the mean number of content problem instances (untransformed data: M = 1.40, SD = 1.22, transformed data: M = 0.86, SD = 0.56) and the mean number of information architecture instances (untransformed data: M = 1.60, SD = 1.51, transformed data: M = 0.98, SD = 0.60). The mean number of content and information architecture problems were not found to be significantly different.

There was also a main effect in the website variable, with different websites having significantly more problem instances than others, $F(2, 44) = 5.29$, $p = 0.009$, $\eta^2_{partial} = 0.194$. Means and standard deviations of problems for each website are presented below.

Table 17. Mean (SD) of participants' problems in each website for untransformed and transformed data.

Website	Untransformed	Transformed
theAA	2.56 (2.07)	1.24 (0.58)
Reed	1.79 (1.43)	0.98 (0.58)
ToysRus	2.25 (1.92)	1.24 (0.70)

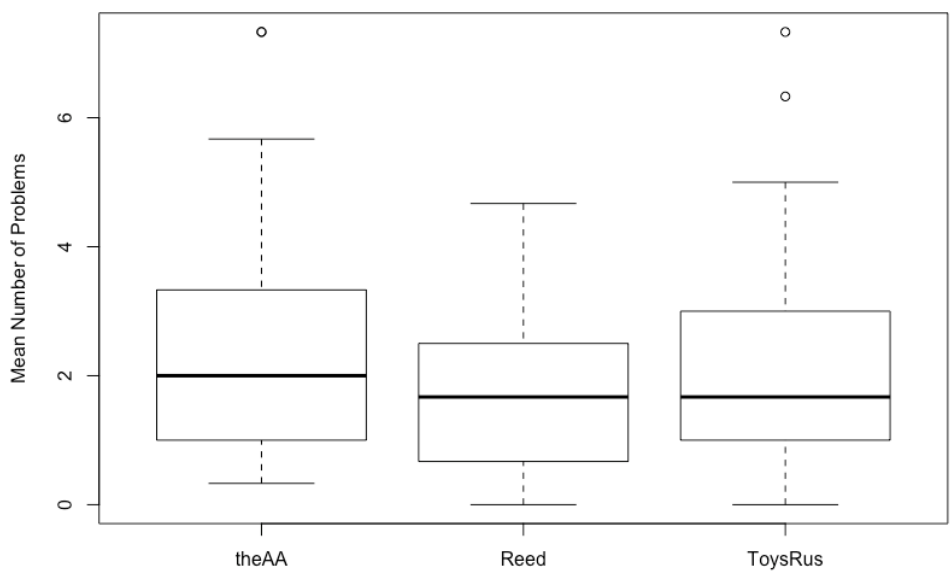

Figure 8. Boxplot showing the distribution of problems per website (untransformed data).

Post-hoc comparisons with Bonferroni correction indicated that the mean number of problem instances encountered on Reed (untransformed data: M = 1.79, SD = 1.43, transformed data: M = 0.98, SD = 0.58) was lower than the mean number of problem instances encounter on theAA (untransformed data: M = 2.56, SD = 2.07, transformed data: M = 1.24, SD = 0.58) and ToysRUs (untransformed data: M = 2.25, SD = 1.92, transformed data: M = 1.24, SD = 0.70). The other comparison, theAA and ToysRUs, was not significantly different in the mean number of problem instances found on them.

From the three-way ANOVA, there was also a significant main effect for the user group, $F(1, 22) = 14.23$, $p = 0.001$, $\eta^2_{partial} = 0.393$. Means and standard deviations of problem instances across websites and problem categories by user group are presented below.

Table 18. Mean (SD) of problems by user group for untransformed and transformed data.

Website	**Untransformed**	**Transformed**
Blind	3.29 (1.59)	1.50 (0.52)
Sighted	1.11 (0.66)	0.80 (0.37)

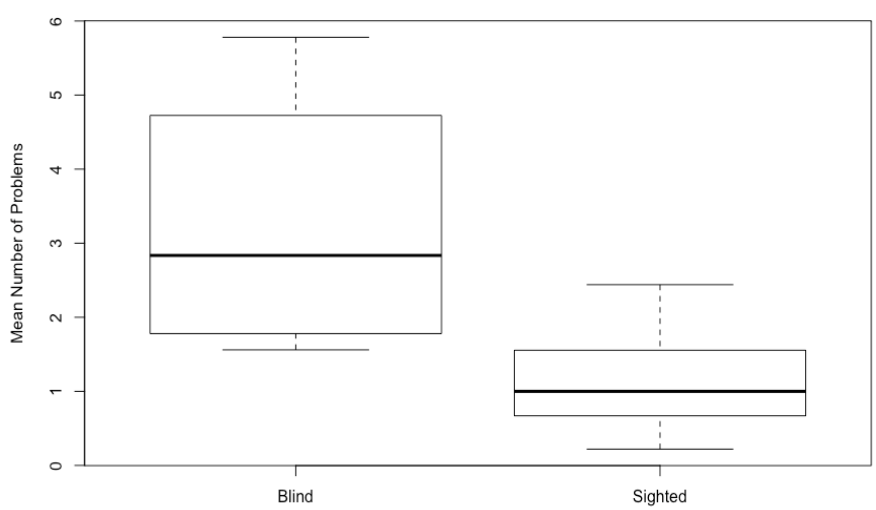

Figure 9. Boxplot showing the distribution of problems per user group (untransformed data).

The mean number of problem instances encountered by blind users (untransformed data: M = 3.29, SD = 1.59, transformed data: M = 1.50, SD = 0.52) was higher than the mean number of problem instances encountered by sighted users (untransformed data: M = 1.11, SD = 0.66, transformed data: M = 0.80, SD = 0.37).

There was also significant interaction between problem categories and user group, $F(2, 44) = 3.59$, $p = 0.036$, $\eta^2_{partial} = 0.140$. Means and standard deviations of problem category instances across websites by user group and problem category are presented below.

Table 19. Mean (SD) of problems in each problem category by user group for untransformed and transformed data.

	Blind		**Sighted**	
	Untransformed	**Transformed**	**Untransformed**	**Transformed**
Content	1.94 (1.41)	1.09 (0.57)	0.86 (0.67)	0.64 (0.47)
Information Architecture	2.39 (1.75)	1.28 (0.59)	0.81 (0.61)	0.67 (0.44)
Interactivity	5.53 (2.81)	2.12 (0.68)	1.67 (1.13)	1.10 (0.49)

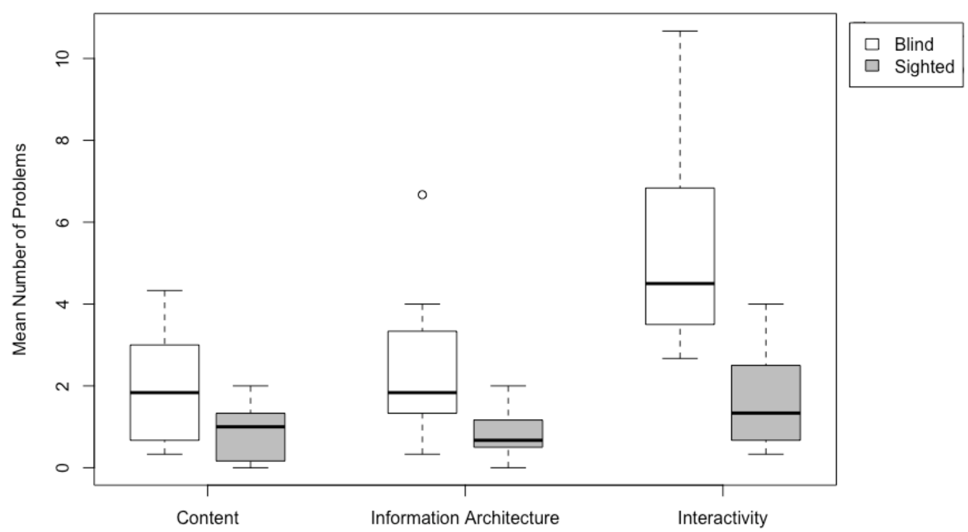

Figure 10. Boxplot showing the distribution of problems per user group and problem category (untransformed data).

An analysis of simple effects shows a significant difference for content category problems between the two user groups, $F(1, 22) = 4.56$, $p = 0.044$, $\eta^2_{partial} = 0.172$. Blind participants encountered significantly more content problem instances (untransformed data: $M = 1.94$, $SD = 1.41$, transformed data: $M = 1.09$, $SD = 0.57$) than sighted participants (untransformed data: $M = 0.86$, $SD = 0.67$, transformed data: $M = 0.64$, $SD = 0.47$).

The analysis revealed that there was also significant difference for information architecture problem category instances between the two user groups, $F(1, 22) = 8.16$, $p = 0.009$, $\eta^2_{partial} = 0.270$. Blind participants encountered significantly more information architecture problem instances (untransformed data: $M = 2.39$, $SD = 1.75$, transformed data: $M = 1.28$, $SD = 0.59$) than sighted participants (untransformed data: $M = 0.81$, $SD = 0.61$, transformed data: $M = 0.67$, $SD = 0.44$).

The analysis also revealed that there was significant difference for interactivity problem category instances between the two user groups, $F(1, 22) = 17.78$, $p < 0.001$, $\eta^2_{partial} = 0.447$. Blind participants encountered significantly more interactivity problem instances (untransformed data: $M = 5.53$, $SD = 2.81$, transformed data: $M = 2.12$, $SD = 0.68$) than sighted participants (untransformed data: $M = 1.67$, $SD = 1.13$, transformed data: $M = 1.10$, $SD = 0.49$).

There was also significant interaction between problem category and website, $F(4, 88) = 9.33$, $p < 0.001$, $\eta^2_{partial} = 0.298$. Means and standard deviations of problem category instances across websites by problem category are presented below.

Table 20. Mean (SD) of problems in each website by problem category for untransformed and transformed data.

	Content		Information Architecture		Interactivity	
	Untransformed / Transformed		Untransformed / Transformed		Untransformed / Transformed	
theAA	0.88 (1.36)	0.63 (0.71)	1.88 (2.52)	1.11 (0.81)	4.92 (4.84)	1.97 (1.04)
Reed	0.67 (0.82)	0.57 (0.60)	1.00 (1.29)	0.71 (0.72)	3.71 (3.10)	1.66 (0.99)
ToysRus	2.67 (2.51)	1.40 (0.86)	1.92 (2.00)	1.11 (0.85)	2.17 (2.46)	1.19 (0.88)

An analysis of the simple effects shows a significant difference for the number of content category problems encountered between the websites, $F(2, 21) = 14.56$, $p < 0.001$, $\eta^2_{partial} = 0.581$. The mean number of content problem instances in ToysRUs (untransformed data: $M = 2.67$, $SD = 2.51$, transformed data: $M = 1.40$, $SD = 0.86$) was higher than in theAA (untransformed data: $M = 0.88$, $SD = 1.36$, transformed data: $M = 0.63$, $SD = 0.71$) and in Reed (untransformed data: $M = 0.67$, $SD = 0.82$, transformed data: $M = 0.57$, $SD = 0.60$). The other comparison, theAA and Reed, was not significantly different in the mean number of content category problems found on them.

The analysis did not reveal any significant differences for the mean number of information architecture problems across the three websites, $F(2, 21) = 3.34$, $p = 0.055$, $\eta^2_{partial} = 0.241$.

However, there was significant difference for interactivity category problems between the three websites, $F(2, 21) = 7.20$, $p = 0.004$, $\eta^2_{partial} = 0.407$. The mean number of interactivity category problems in theAA (untransformed data: M = 4.92, SD = 4.84, transformed data: M = 1.97, SD = 1.04) was higher than in the ToysRUs (untransformed data: M = 2.17, SD = 2.46, transformed data: M = 1.19, SD = 0.88). Also, the mean number of interactivity category problems in Reed (untransformed data: M = 3.71, SD = 3.10, transformed data: M = 1.66, SD = 0.99) was higher than in the ToysRUs (untransformed data: M = 2.17, SD = 2.46, transformed data: M = 1.19, SD = 0.88). The other comparison, theAA and Reed, was not significantly different in the mean number interactivity category problems found on them.

There was no significant interaction between website and user group, $F(2, 44) = 1.04$, $p = 0.361$, $\eta^2_{partial} = 0.045$. Means and standard deviations of problems per website by user group are presented below.

Table 21. Mean (SD) of problems in each website by user group for untransformed and transformed data.

	Blind		**Sighted**	
	Untransformed	**Transformed**	**Untransformed**	**Transformed**
theAA	4.03 (1.86)	1.66 (0.42)	1.03 (0.64)	0.82 (0.37)
Reed	2.50 (1.51)	1.27 (0.59)	1.08 (0.94)	0.69 (0.41)
Toysrus	3.28 (2.21)	1.57 (0.75)	1.22 (0.72)	0.90 (0.49)

There was no three-way interaction between website, problem category and user group, $F(4, 88) = 1.79$, $p = 0.139$, $\eta^2_{partial} = 0.075$.

The analysis revealed some interesting differences regarding the number of problem instances that bear a further investigation. First, there is a difference between the websites regarding the number of problems found by users, as well as differences regarding the category of problems found. On the one hand, this is quite positive, given that it was wanted a robust collection of problems from the websites. However, it is important to examine if these differences have generated problem instances that are unusual in some way or highlight particular aspects of websites that generated more problems. Secondly, there are definite differences between the user groups that are being highlighted within the data set that are of interest. Specifically, blind participants found more problem instances across all the three problem categories that had instances by both user groups. To further understand these differences a further analysis of the problem differences between user groups was conducted.

5.3.3. Analysis of problem differences between user groups

The analysis of the participants' problems revealed some interesting differences in the problem categories encountered on the web between the two user groups. There were problem categories distinct to each user group, but also problem categories that had instances by both users but blind users encountered significantly more problems. To further understand these differences, an analysis of the problem types of each problem category was performed.

Physical Presentation Problems

Physical presentation problem category involved issues related to the physical presentation of the page content. Without question, problems of this category were distinct to sighted users as they were about the visual presentation of the page. Table 22 shows the frequency of each physical presentation problem type.

Table 22. Physical Presentation problem types with their frequency.

Problem type	Frequency
Text/ interactive elements not large/clear /distinct enough	11
Page layout unclear/confusing	10
Key content/ interactive elements, changes to these not noticed	2
“Look and feel” not consistent	2

Looking into the characteristics of t*ext/ interactive elements not large/clear /distinct enough* problem type, most of the problems reported were about the font size of the page being very small. Figure 11 shows an example in theAA. Users felt that the font size of the menu items was very small which made it more difficult to read it.

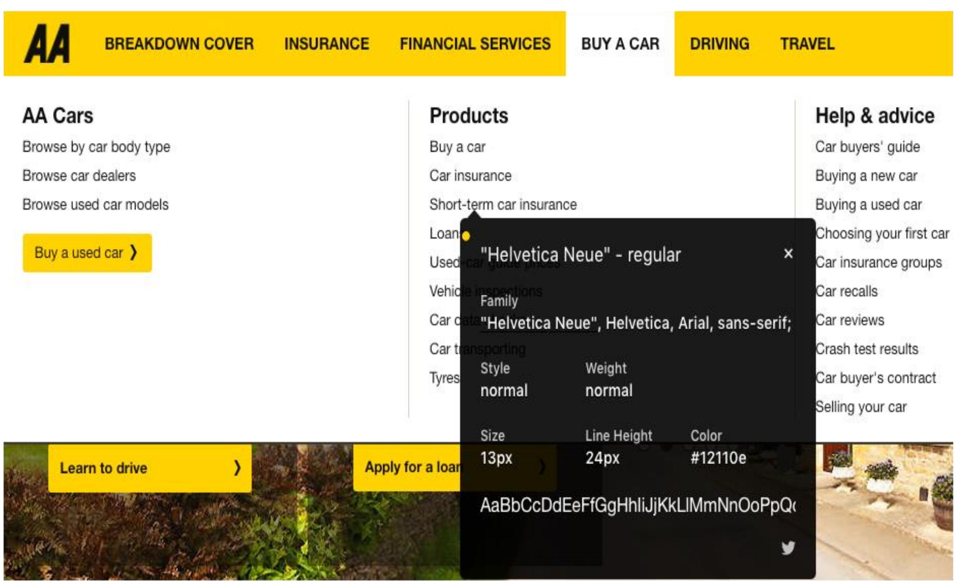

Figure 11. Example of a text/ interactive elements not large/clear /distinct enough physical presentation problem type.

Characteristics of *page layout unclear/confusing* problem type involved issues with the presentation of the content causing difficulties to sighted users to use the page. For example, users encountered difficulties using the filtering options on the page as they could not see all the filtering options and had to constantly scroll up and down to select all the required filtering options.

When considering the *"look and feel" not consistent* problem type, this involved issues with the visual presentation of the page not being consistent across the different pages. Figure 12 and Figure 13 shows an example of this problem. The homepage (Figure 12) had different visual presentation compared to the buy used cars page (Figure 13) in theAA.

Figure 12. Example of "look and feel" not consistent in theAA homepage.

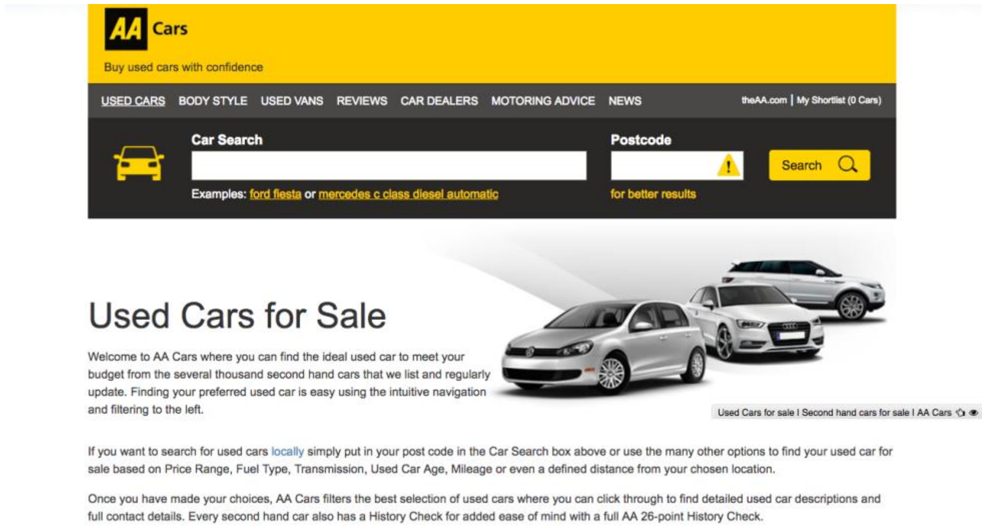

Figure 13. Example of "look and feel" not consistent in theAA buy used cars page.

Technology problems

Technology problem category included issues related to the browser users were using and the assistive technology. Table 23 presents the frequency of the technology problem types.

Table 23. Technology problem types with their frequency.

Problem type	Frequency
Browser	5
Assistive Technology	21

All the issues reported for the *browser* problem type were about the users getting stuck at the browser toolbar and not being able to access the page content. Characteristics of *assistive technology* problem type largely included the screen reader behaving unexpectedly. For example, the screen reader focus was jumping to a random section on the page when the page refreshed instead at the top of the page, reading the page title multiple times, losing focus from the page, not perceiving correctly the page content and pronouncing incorrectly the page content. An example, where screen reader pronounced incorrectly the page content was in ToysRUs. A product contained the word “Toucan”, but the screen reader read it out as “Token”.

Figure 14 shows an example where screen reader did not perceive the page content correctly. The VoiceOver rotor, functionality to navigate quickly to elements on the page, such as links, headings, form elements, tables, was showing that there were not any available form inputs on the page. The rotor was incorrectly showing that there were not any form elements on the page when there were.

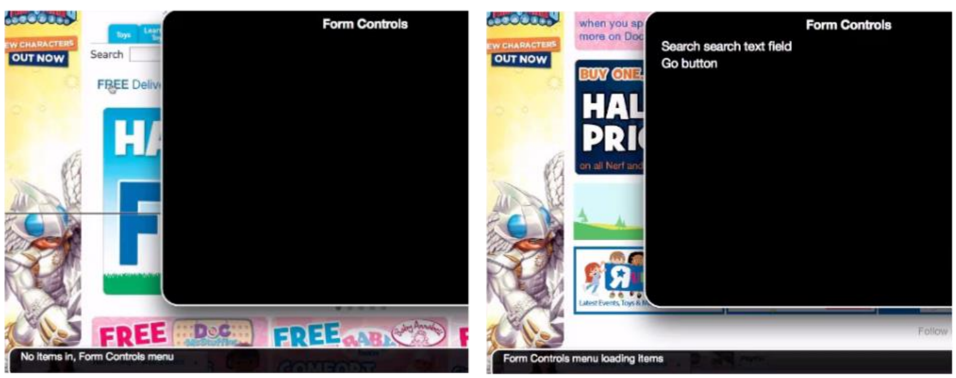

Figure 14. Example of a characteristic of a technology problem type.

Content problems

Content problem category is comprised of six content problem types. Table 24 shows the frequency of each content problem type by user group.

Table 24. Frequency of content problem types by user group.

	User group	
Problem Type	**Blind**	**Sighted**
Too much content	9	5
Content not clear enough	15	7
Content not detailed enough	9	2
Content inappropriate or not relevant	34	15
Terms not defined	-	1
Duplicated or contradictory content	3	1

As can be seen from the table (above), almost both user groups had instances of each content problem type. The only content problem type that was distinct to sighted users was *terms not defined.* This problem type involved issues with users encountering terms that they did not know and could not find out more information.

Even though the two user groups encountered similar content problem types, the number of distinct problems (i.e. problems that were matched as being the same) where both user groups encountered them were very low. Figure 15 shows the distribution of the distinct content problems between the two user groups. As can be seen from the figure, 70% (N = 41) of the problems were encountered only by blind users. Problems that were distinct to sighted users accounted for 22% (N = 13). The overlap of the distinct problems between the two user groups was only 8% (N = 5). Even though the two user groups encounter a similar type of content problems when it comes to the distinct problems the overlap between the two user groups is very low.

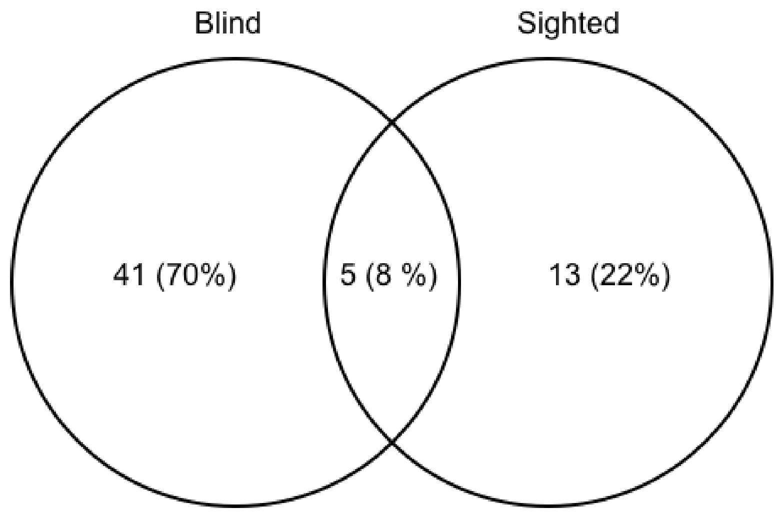

Figure 15. Numbers and percentages of the distinct content problems identified for the two user groups.

To further understand what problems blind users are finding that are different from sighted users, the problems were examined for their causes or common characteristic that explain the differences. When considering the content problem type *too much content*, all the problems reported by sighted users were about the pages themselves having too much information. For example, in Reed homepage, users felt there was too much information to take in. As one sighted participant put it:

> *"The website seems to be so long. I just wanted to make sure I am not missing anything. There was a bit too much information to take in…I just felt a bit overwhelmed, a bit lost in information there" (P17).*

In comparison, when looking at the problems relating to *too much content* reported by blind users, the problems are not related only to the whole page but instead focus on specific sections of the page where they were overwhelmed. Figure 16 shows an example where blind users reported issues at the level of specific sections of the website. Some blind users felt that the list of cars in theAA had too much information about each car listed.

Figure 16. Example of too much content problem type by blind users in theAA.

Looking into the characteristics of *content not clear enough* problem type, sighted users largely described issues with the visual clarity of the images. For example, some users reported issues with some images of scooters in ToysRUs not been very clear that there were scooters. For blind users, almost all of the problems, 14 out of 15, were about the lack of alternative text on content, which is a well-known issue. Much of this comes from not meeting existing standards of providing an alt attribute for an image, which results in the users hearing the file name through the screen reader. Figure 17, shows an example of this problem type. The screen reader is reading the image as “88771222214686.gif”, meaning blind users cannot identify the content of the image.

Figure 17. Example of content not clear enough reported by blind users in ToysRUs.

The most frequent content problem type for both user groups was *content inappropriate or not relevant.* Four out of five shared distinct problems between the two user groups was of this problem type. Both user groups reported issues for information not seem relevant to the task. For example, the list of results contained results that did not match with the users' selections. The comments below present participants' comments when the page was showing scooters that were not in the age range that users specified:

"I have filtered for 5-7 year old, and the product I was looking at was not in that age range". (P3, blind)

"the age group of the scooter in the results is not consistent with the filtering I selected." (P18, sighted).

Users in both user groups reported issues with the advertisements on the pages, however, they were largely different problems. The reason for this seems to relate to the position of the advertisements in areas of the page. Blind users reported issues with advertisements as they were positioned in such a way that they interrupted the task as they progressed through the page. In Figure 18, an example of this is presented, where the screen reader user would proceed through the filtering options and then encounter an advertisement unexpectedly. Users were not expecting to find an advertisement between the filtering section and the results section of the page. In comparison, the problems reported by sighted users about the advertisements on the page were about the page having too many advertisements rather than particular sections of the page. Both problems fail under the same problem type but are different.

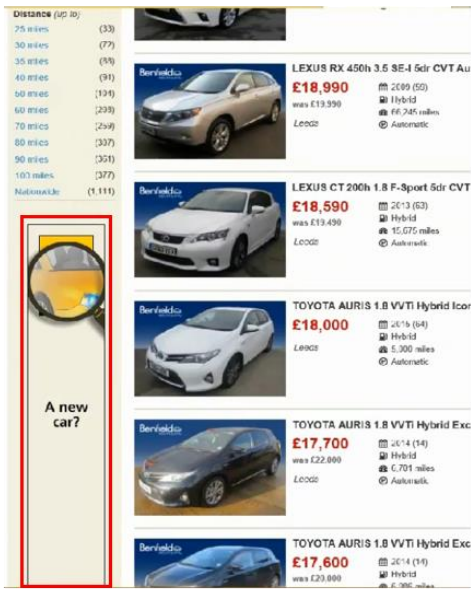

Figure 18. Example of irrelevant content (advertisement) reported by blind users in theAA.

A very distinct example of *content inappropriate or not relevant* problem type for blind users was the references to coding language as part of the page content. For instance, blind users were reading about JavaScript as part of the page content, where visually there was nothing. Blind users perceive this as part of the page content and were confused as to what that has to do with their task.

Information Architecture problems

Information architecture category is another category that had problem instances by both user groups. This category is comprised of six information architecture problem types. The frequency of each information architecture problem type is presented in Table 25.

Table 25. Frequency of information architecture problem types by user group.

	User group	
Problem Type	**Blind**	**Sighted**
Content not in appropriate order	21	22
Not enough structure to the content	23	-
Structure not clear enough	22	-
Headings/titles unclear/confusing	13	1
Purpose of the structures not clear	1	-
Content not where the user expected it to be (between or within page)	6	6

As presented in Table 25, half of the information architecture problem types were distinct to blind users. This included *not enough structure to the content*, *structure not clear enough* and *purpose of the structures not clear.*

The difference of the information architecture problems is not only presented in the problem types but also in the distinct problems between the two user groups. Figure 19 shows the distribution of the distinct information architecture problems between the two user groups. As can be seen from the figure, 82% (N = 55) of the problems were encountered only by blind users. The problems that were encountered only by sighted users accounted for 10% (*N* = 7). The overlap of the distinct problems between the two user groups was only 8% (*N* = 5).

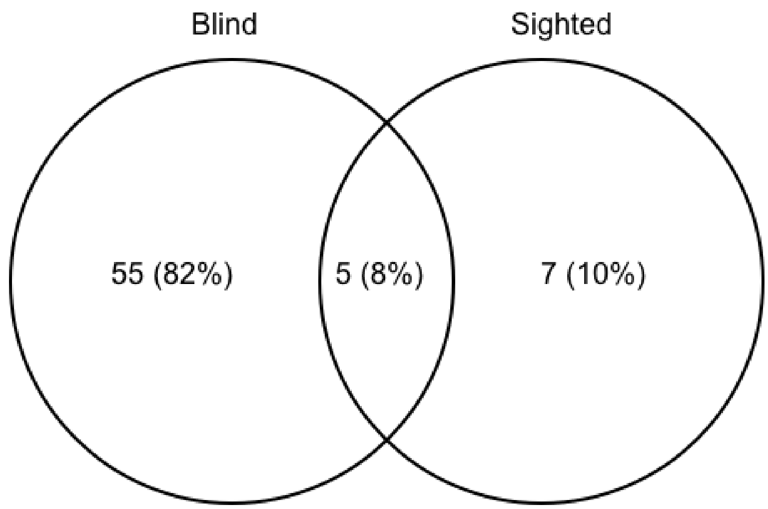

Figure 19. Numbers and percentages of the distinct information architecture problems identified for the two user groups.

To further understand the differences of the information architecture problems between the two user groups an investigation into the characteristics of the problem types that were encounter only by blind users but also the problem types that had instances by both user groups was conducted.

Looking into the problem types that were distinct to blind users, these problems were involved with the page structure being difficult to navigate and finding particular information. For the not *enough structure to the content* problem type, two-third of the characteristics of the problems were about the lack of headings to structure the page content. Blind participants reported issues of lack of headings to structure the page content overall, but also specific areas of the page such as the results sections, the filters section and each result in the list of results. Figure 20 shows an example of the lack of headings to structure specific areas of the page content. Blind participants expected the list of results to be structured using headings that included appropriate markup, which was not present, whereas sighted users see a visually distinct heading.

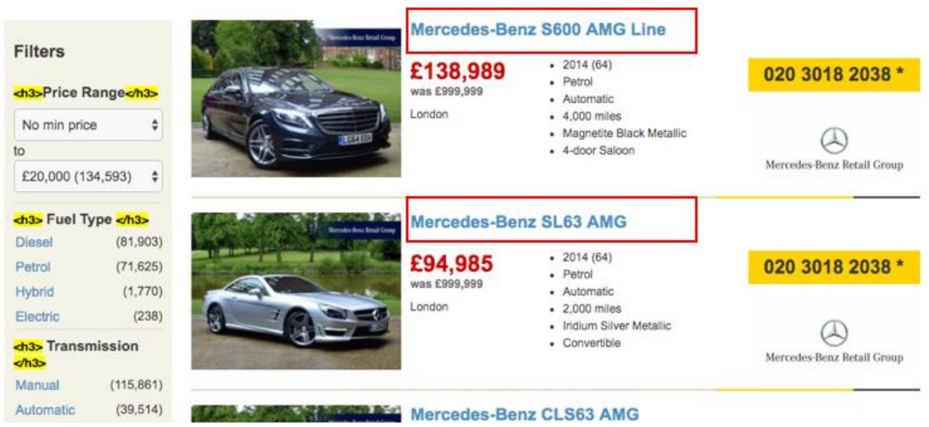

Figure 20. No headings used to structure the results list (theAA).

Other characteristics of this problem type involved heading levels not being in hierarchical order and tables not being well structured. The heading levels not being in hierarchical order means that the headings do not follow a logical order with missing heading levels. Figure 21 shows an example of headings not being in hierarchical order in Reed. There were more than one headings of level one, as well as missing heading levels, meaning going from heading level 1 to heading level 3. Regarding the tables not being well structured, participants reported having difficulties navigating through the table content.

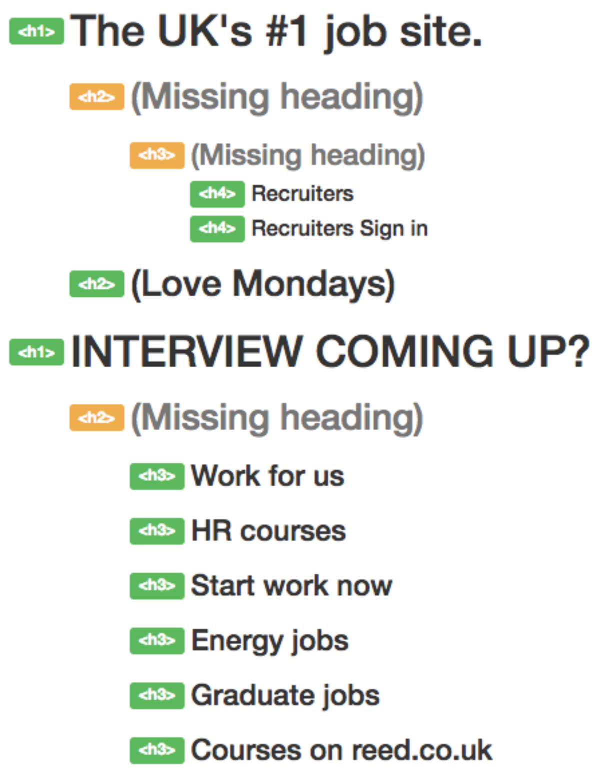

Figure 21. Headings structure does not follow a hierarchical order (Reed).

For the *structure not clear enough*, most of the characteristics of the problems reported by blind users were about getting lost on the page, not being able to find particular information, not easy to navigate and finding particular information on the page. As some users put it:

“It was not easy to navigate to the products...It also made me think did I miss the thing” (P10, ToysRUs)

“I have no idea where I am, on that page, at all…I started realising I was way down at the bottom of the page again, and I did not know why…still could not find the scooters listed” (P11, ToysRUs)

"I could not quite figure out the page structure…it was not easy to access the information I wanted. It just needs a simpler structure." (P5, theAA)

Looking into the information architecture problem types that had instances by both user groups, for the *headings/titles unclear/confusing* blind users encounter much more problem instances than sighted users. Both user groups reported issues with headings not being descriptive. Blind users also reported issues with headings, providing too much description, not finding the headings useful and headings been confusing. Figure 22 shows an example where the headings of the page were perceived not being helpful. There were duplicate headings in the list of headings, which did not provide any helpful information to users "Not sold in store".

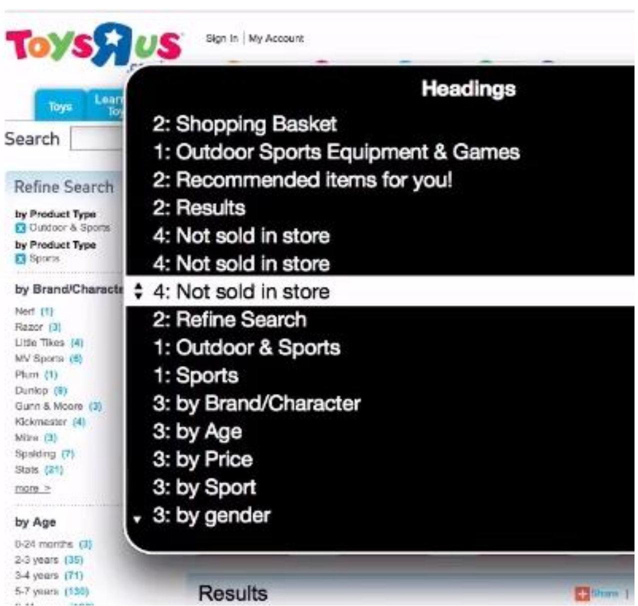

Figure 22. Heading not descriptive (ToysRUs).

For the c*ontent not in appropriate order* problems, both user groups reported problems with the order that the page content was presented, with relevant task content not prioritised. This problem type was the one with the most distinct problems that encountered by both user groups, with three out of the five shared distinct problems. Examples included important information being at the bottom of the page, content not organised in alphabetical order and content not prioritised. An example of a problem that both user groups reported was the order that encountered the filtering options was not prioritised. As some participants said:

> *"I think some of the formatting here. The prioritise of the options. I had to click through the brands of bikes and scooters before I could then select the age. I think the age should be further up in your filter search and the brands should go below." (P5, blind, ToysRUs)*

> *"The brands have been given too much priority…Brands are important, but the age of your child will be more important to make decision." (P13, sighted, ToysRUs)*

There were many problems distinct to blind users. For example, in ToysRUs participants had to search for a scooter, but in the navigation menu participants were first finding the option of outdoors and sports and then the bikes and scooters. As blind users do not have a full overview of all the options that are available on the navigation instantly, some participants found first the outdoors and sports category and thought that as you ride your scooter outside, that it should be the correct category to look at, without looking to the rest of the options available. After realising that they were looking in the wrong category and there was another category specifically for bikes and scooters, they reported that they could be avoided spending all of this time looking in the wrong category if the options were listed in more appropriate order. In contrast, sighted users could easily find the correct category to look as they had an instant overview of all the navigation options.

Interactivity problems

Interactivity problems was the category with the largest number of problems found by both user groups. This problem category is comprised of 12 problem types. Table 26 shows the frequency of each interactivity problem type by user group.

Table 26. Frequency of interactivity problem types by user group.

	User group	
Problem type	**Blind**	**Sighted**
Lack of information on how to proceed and why things are happening	2	4
Labels/instructions/icons on interactive elements not clear	38	28
Duplication/excessive effort required by user	48	3
Input and input formats unclear	9	4
Lack of feedback on user actions and system progress	31	-
Options not logical/complete	2	1
Too many options	6	3
Interaction not as expected	39	12
Interactive functionality expected is missing	5	2

Interactive and non-interactive elements not clearly identified	11	1
No labels at interactive elements	4	2
No consistency between interactive elements	4	-

As presented in Table 26, both user groups encounter instances from almost all interactivity problem types except two interactivity problem types that were distinct to blind users. These were the *lack of feedback on user actions and system progress* and the *no consistency between interactive elements.* The lack of feedback on users' actions was one of the most frequently reported problems by blind users. This problem type involved issues with blind participants not being able to relate what was happening on the page in relation to their actions. Most of the problems reported by blind users were when users interacted with the filtering options on the websites, 29 out of 31 feedback problem types. Each time users selected a filtering option there was no feedback that their action was performed successfully. The websites were refreshed and updated their content, however, blind users were unsure whether their action was performed successfully since they did not receive any form of feedback.

The overlap of distinct problems between the two user groups was very low. Figure 23 shows the distribution of the distinct interactivity problems between the two user groups. As can be seen from the figure 71% (N = 108) of the distinct problems were encountered only by blind users. Sighted users encountered 21% (N = 33) of the distinct problems. The overlap of the distinct problems between the two user groups was only 8% (N = 12). To further understand what interactivity problems blind users are finding that are different from sighted users, the characteristics of the problems were examined to explain the differences.

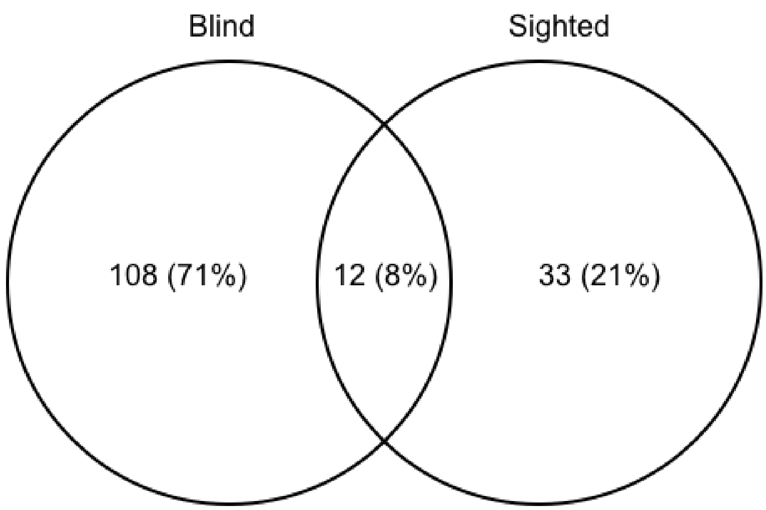

Figure 23. Numbers and percentages of the distinct content problems identified for the two user groups.

Both user groups had instances of *labels/instructions/icons on interactive elements not clear* problems. Examples of problems included links not being descriptive enough, similar links on the page and labels at interactive elements not being clear. Blind users also reported issues with image links not having an accessible description and links having too much information. Figure 24 shows an example where blind users

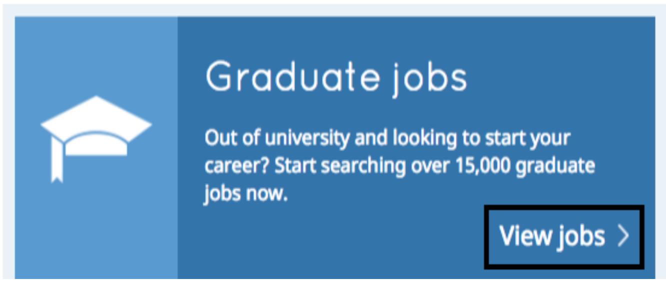

Figure 24. View jobs link not descriptive (Reed).

perceived a link not being very descriptive. The link description was "view jobs". Blind users were unsure whether the link was for graduate jobs or general jobs, as it was not clear from the link description itself. In contrast, sighted users could relate the link with the heading that was before the link, that it was referring to graduate jobs.

Another example included the link to clear the selected options in the filtering section. The link description was "clear" (theAA) and "remove" (ToysRUs). However, blind users had difficulties understanding what action will be performed. Sighted users did not experience any problems with the links, as it was clear what action will be performed as the link was positioned next to the filtering option.

There were also problems with the numbers next to each filtering option in the filtering section. Figure 25 shows an example where a number is next to each filtering option. The purpose of the number is to inform users how many results will be returned for that option. Sighted users did not encounter any problems understanding what the numbers were referring to in the list. In contrast, blind users reported problems understanding what the numbers are about, as there was no context of what the number means as the number was not linked to the field in the code in any way.

Figure 25. Numbers next to each filtering option indicating how may results will be returned.

The *duplication/excessive effort required by user* problems were the most frequently reported problem by blind users across all problem types. Most problems reported by blind participants were about a specific design feature in relation to the filtering browsing of page content. Each time users selected a filtering option the page updated its content. However, the screen reader focus was going at the top of the page and each time users had to retrace their steps to where they were interacting with the screen reader to progress with their task. In contrast, sighted participants only reported issues with having to enter information twice or some tasks required too many steps to be completed. As presented, there were many differences in the characteristics of this interactivity problem between the two user groups.

The *interaction not as expected* problems were encountered more often by blind users. For blind users, this mainly involved issues with links not working and interacting with the page in a way that was unexpected. An example by blind users included participants trying to interact with the skip link of the page, but the link was not working. Sighted users also encounter issues with interacting with the page in a way that was unexpected. For example, some sighted participants had difficulties to use the menu. When users selected a menu option, the website was taking the users to that page. However, users expected for the website to show the submenu options of the selected menu option, rather than taking them to a different page.

When considering the *interactive and non-interactive elements not clearly identified* problems, blind users encountered much more problems than sighted users. The characteristics of this problems largely differ between the two user groups. Sighted users reported having problems trying to interact with images that were not interactive. In contrast, blind users had difficulties perceive the interactive elements of the page. Figure 26 shows an example in theAA where the filtering options of the page are not read as interactive elements via the screen reader. This problem is caused because the page is using not the appropriate markup language for interactive elements, which is causing screen readers perceiving them as non-interactive elements.

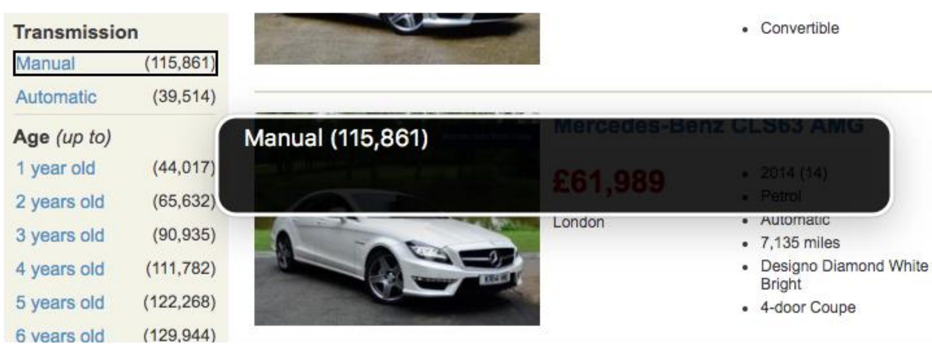

Figure 26. Interactive elements not properly implemented (theAA).

Looking into the *no labels at interactive elements* problems, two different set of problems were found. The first set of problems involved issues for both user groups reporting problems with interactive elements not having labels. For example, in the filtering section in Reed, there were two inputs for the salary range that did not have a label.

Another set of problems that were distinct to blind users were interactive elements having a label next to them, but the label was not programmatically associated with the interactive element. For example, in ToysRUs, there were two dropdowns to sort the results of the page and to select how many results to show per page. The dropdowns had a label next to them, which visually was making it clear what the elements were referring to. However, the labels were not programmatically associated with the dropdowns. Blind users had difficulties understanding what the interactive elements were about as there was no information associated with them.

5.3.4. Problems' Severity Ratings

The severity of the problems users encountered was also collected. Table 27 presents the mean severity rating for each problem type by user group.

Table 27. Mean (SD) of problems' severity rating for each problem type by user group. The problem types where SD is not reported (N/A) is because there was only one instance of that problem type.

Problem Type	**Blind**	**Sighted**
Physical Presentation		
Text/ interactive elements not large/clear /distinct enough	-	3.45 (0.93)

Page layout unclear/confusing	-	3.00 (0.94)
Key content/ interactive elements, changes to these not noticed	-	2.50 (0.71)
"Look and feel" not consistent	-	3.00 (1.41)
Technology		
Browser	3.60 (1.52)	-
Assistive Technology	3.00 (1.34)	-
Content		
Too much content	3.44 (1.42)	2.80 (0.45)
Content not clear enough	2.73 (1.10)	2.57 (1.40)
Content not detailed enough	2.89 (1.17)	3.50 (0.71)
Content inappropriate or not relevant	2.71 (1.19)	2.80 (0.94)
Terms not defined	-	2.00 (N/A)
Duplicated or contradictory content	2.00 (1.00)	3.00 (N/A)
Information Architecture		
Content not in appropriate order	2.71 (1.06)	3.05 (1.00)
Not enough structure to the content	3.00 (1.13)	-
Structure not clear enough	3.55 (1.41)	-
Headings/titles unclear/confusing	2.77 (1.01)	4.00 (N/A)
Purpose of the structures not clear	4.00 (N/A)	-
Content not where the user expected it to be (between or within page)	2.83 (0.75)	2.50 (0.55)
Interactivity		
Lack of information on how to proceed and why things are happening	2.50 (0.71)	3.00 (0.82)
Labels/instructions/icons on interactive elements not clear	2.68 (1.40)	2.68 (1.16)
Duplication/excessive effort required by user	2.92 (1.01)	4.33 (1.15)
Input and input formats unclear	2.00 (0.87)	2.50 (1.00)

Lack of feedback on user actions and system progress	3.52 (1.03)	-
Options not logical/complete	2.00 (0.00)	2.00 (N/A)
Too many options	3.67 (1.75)	3.67 (0.58)
Interaction not as expected	3.15 (1.16)	2.67 (1.07)
Interactive functionality expected is missing	2.60 (0.89)	3.50 (0.71)
Interactive and non-interactive elements not clearly identified	3.00 (1.48)	2.00 (N/A)
No labels at interactive elements	1.50 (1.00)	3.50 (0.71)
No consistency between interactive elements	3.00 (1.83)	-

As it is shown in the subsection 5.3.3, the problems the two user groups encounter largely differ. Thus, an analysis of the severity rating of problems between the two user groups was not performed, as it would not provide any valuable information for understanding the problem differences between the two user groups.

5.4. Discussion

This study set out to investigate the similarities and differences of the problems encountered on the web between blind and sighted users. This is the first study, to our knowledge, to examine the differences of the problem encountered between the two user groups on the web.

Users' experience on the web was also investigated and showed that blind participants are still in disadvantage compared to sighted participants regarding the effective use of websites. The result of the task success rates for blind users (53%) is broadly consistent with the findings of previous research (Disability Rights Commission, 2004; André Pimenta Freire, 2012; Petrie & Kheir, 2007), with task success rates of 53%, 55.96% and 50.7%, respectively.

The results of the study indicate that there are sets of problems that are distinct to each user group. Starting from the main category of problems, it was found that technology category problems are distinct to blind problems. Most of these problems reported were about the assistive technology interaction behaving in an unexpected way. Problems with the technology interaction blind users were using are reported since the early stages of the web (Oppenheim & Selby, 1999). Even though web technologies move forward, there are still technology mismatches between the screen readers, the web code and the browsers.

Looking into the problem types, there were a few problem types that were distinct to blind users. This mainly included problems with the structure of the page and the lack of feedback on users' actions.

Blind participants encountered many problems with the structure of the page being difficult to navigate and find particular information. One of the most frequently reported problems was the lack of headings to structure the page content. The use of headings to navigate through the page content is one of the most commonly techniques blind users rely on (Power et al., 2013; WebAIM, 2014) and can play an important role for blind users' experience on the web, as it can help them navigate through the page content quickly (T. Watanabe, 2009) and get an overview of the overall structure of the page content (Power et al., 2013). If the page is not using headings for content structure or if the structure of the headings on the website is poor blind users will have difficulties navigating and finding information as they will have to navigate using different techniques, for example, navigate sequentially through the page content. Then, they may encounter other problems, such as irrelevant content, which it can add to their mental load as they try to relate it to the rest of the page content. The problems with the page structure corroborate the findings of previous work in the field about the problems around the page structure (Disability Rights Commission, 2004; Power et al., 2012). Interestingly, the structure that was missing in the code of the page was largely there for sighted users. Many of the websites visually indicated the main page elements with either bolder text or bigger font size. Although visually it was easy to locate the different sections of the page, blind users did not have the same experience.

Another problem that was distinct to blind users was the lack of feedback on users' actions. This problem type was also found in the previous study of this thesis (Chapter 4) but in much less frequency. Talking about this issue, some participants said:

> *"I found it difficult to filter the search. It was not easy…when I clicked on them, I found that it did not make any difference on the search. There was no feedback whether it accepted my selection." (P5, Reed)*
>
> *"When I selected the hybrid, I can hear the page reloading, that means it accepted the criteria. I had to assume that it saved it. There was no feedback whether it was selected or not." (P11, theAA)*

Users were not sure whether their action was performed or not. Similar problems reported in other studies in the literature (R. Babu & Singh, 2009, 2013b; Coursaris et al., 2014; André Pimenta Freire, 2012; Giraud et al., 2011). For example, André Pimenta Freire (2012) noted: "users performed an action on the website and could not identify any feedback that the action had been performed". Coursaris et al. (2014) noted: "when a food item was added, there was no notification; hence, user were often confused about whether or not the item had been added". In the present study, most of the feedback issues reported were in relation to the search and filtering browsing of content. Each time users selected a filtering option the page updated its content and refreshed. However, it was not clear for screen reader users whether their action was performed successfully or not. The lack of feedback can disorient users and lead to a cognitive strain, as they will have to compare the website model they have in working memory with that they are finding in the page. Participants task time will increase as they will try to use different tactics to identify what is happening on the page (Vigo & Harper, 2013).

There were also many problem types that had instances by both user groups. However, when the problems were closely looked, it was found that the characteristics of the problems reported, largely differed. An example of the content problems was that both user groups had issues with the content of the page not being clear. However, the problem causes were completely different. Sighted users had problems with the clarity of visual information conveyed through images. The problems reported by blind users mainly involved the lack of alternative text on images, which resulted in not being able to decode the meaning of images as there was not any information about them. Issues with the lack of alternative text of images is a well-known problem on the web that still seems to exist. Providing alternative text on images can be difficult sometimes as it is at the discretion of designers whether an image should be marked as decorative or whether it conveys information. Example paradigms of appropriate use of alternative text on images are available (Reid & Snow-Weaver, 2008; WebAIM, 2017) which can help to address this problem.

Another problem that both user groups had was the information overload. A further look into these problems shows that users reported the problems at different sections of the website. All problems reported by sighted users were about the information on the page having too much content overall. On the other hand, blind users also had issues with specific sections of the page having too much content. This problem seems to relate to how blind users try to understand how the page content is structured (Power et al., 2013). Blind users often use the landmarks of the page (e.g. headings) to understand the structure of the page then they explore the content around it. When they tried to interact with specific areas of the page if the area is overloaded with information they can get hindered. This difference highlights the importance of designers giving more attention to the amount of information provided in each section of the page.

Irrelevant information was another problem that had instances by both user groups. There were many distinct examples of this problem type to blind users. For example, some of the problems reported by blind users were:

> *"The advert between filters and results it distracts you. The position of the advert is getting into your way." (blind, theAA, P12)*
>
> *"There is a flash right in the middle I guess, but kind of between the top where is your search.. I think it is under where I selected the prices and the filters. There is a bit of flash that talks about loans…it kind of gets into the way." (blind, theAA, P7)*
>
> *"There are some companies-advertisements before the jobs. That is irrelevant." (blind, Reed, P9)*
>
> *"When I am clear what I am looking for and this only shows me products for my child…It annoyed me that I had to sit through text*

that it was just telling that toys are good for kids, kids like toys…that further slow me down..instead of getting on with finding what I wanted..” (blind, ToysRUs, P10)

“I did not like finding references to JavaScript. It is completely irrelevant to me. I do not know that JavaScript. I know it is interesting to web designer people, it is completely irrelevant to me and another thing getting on my way.” (blind, ToysRUs, P1)

Many of the problems reported by blind users relate to the navigation differences between the two user groups. When blind users navigate through the page content sequentially, they cannot skip information that is irrelevant to their content. Moreover, this can be very problematic in the case when the page content is not properly structured with headings, as users will not be able to use the headings to skim through the page content. On the contrary, sighted users can have a full overview of the page content instantly and can skim through the page content quickly. Content that seems irrelevant to their task can be skipped. An example to further support this was the problem that encountered by both user groups with the advertisements on the page. Sighted users reported issues that the pages were having advertisements overall, whereas blind users perceived the adverts interrupting their navigation flow as they were positioned between the filtering and results section.

A problem type that was frequently reported by blind participants was the duplication and excessive effort required to perform the tasks. Most of the problems reported by blind users were in relation to the filtering browsing of page content. As some participants said:

“Again I ordered the results by price and it stacked way back to the top of the page again…I had to navigate to the list” (blind, theAA, P3)

“Each time I apply a filter it reloads the page and I have to scroll all the way back to the filters and set the filters again. It does not allow me to set all the filters together and then I can search. For each option, it does a reload which just adds the amount of time and plus a bit more.” (blind, ToysRUs, P2)

“I selected engineering and again it took me to the main navigation bar full of links that you do not want. You start again going through all the navigation bar. But you do not want that.” (blind, Reed, P5)

Sighted users also had instances of this problem type. However, the nature of the problems reported by sighted users was very different. For example, one sighted participant said:

“I selected the graduate’s link but the form resets. I had to enter all the filtering criteria again.” (sighted, Reed, P15)

Blind users' problems were mainly about the extra effort added due to the refreshing of the page, which sighted users did not have. Blind users had to retrace their steps to progress with their task each time the page refreshed and had to go through all the task requirements that sighted users had to do.

Most of the problems reported by blind users were around the interaction with the filtering options of the websites. The websites refreshed and updated their content. That was causing the user focus at the top of the page each time an update on the page was happening, and users had to retrace their steps back to progress with their task. A possible design solution to address the excessive effort of blind users when interacted with the filtering option is to update the page content without the page refreshing. However, a drawback of this approach is the lack of feedback on users' actions as users will not know whether their action was performed or not. An alternative approach is websites first to allow users to select all filtering options and then users can request a change of context through the selection of a submit button. Thus, the refreshing of the page will happen once. This approach can highly benefit users when selecting multiple filtering options. However, neither website in the present study provided a submit button for initiating a change of context, as the update of the page content was happening as users were selecting a filtering option.

The findings of the present study provide additional support that the problems encountered by the two user groups differ (Petrie & Kheir, 2007) but also expands our current understanding of the problem differences. The differences in the characteristics of the problems encountered on the web between the two user groups highlight the importance of considering the diversity of users' needs when designing websites. As Horton and Quesenbery (2014) mentioned "you have to know the people you are designing for. And that includes people with disabilities". Websites should be designed to accommodate the needs of different user groups, and that includes blind users. The results of the study demonstrated that the problems blind and sighted users encounter on the web largely differ, which can imply that fixing the problems sighted users have does not necessarily means the problems blind users have will be addressed.

A large number of problems encountered by blind users may influence each other and impact the usability of websites. Problems with the page structure can make it difficult for users to retrace their steps to where they were interacting with the page content when the page refreshes. The lack of feedback can leave users disoriented whether a change of context happened. Some of the issues had to do with how page information is organised and presented on websites. Problems with irrelevant and overwhelming content were mainly isolated to specific areas of the page. These problems are going to be exclusive to the website and will be tied deeply to the understanding of the content writer of how blind people will interact with the page content. For blind users to find the content, it presupposes that users had to traverse the page using their screen reader. However, it is unclear whether the content itself was reducing the users' experience.

During the traversal, blind users will have to maintain their understanding of where they are on the page due to the poor page structure, and they will also try to interpret the results of their actions due to the lack of feedback and all at the same time they trying to build and maintain the overall overview of the page. Participants cognitive load will probably be very high as they try to do all these things, and that may result even moderately complex content to seem overwhelming. Further, users will try to interpret the state of the page that lacks sufficient feedback, when they encounter content that seems irrelevant it may be because they cannot relate it into their understanding of what the website is doing. Clearly, content like advertisements is irrelevant to the users' task. However, because of the number of problems encountered before they actually find irrelevant content, blind users may consider some items irrelevant that are in fact meaningful. For example, some users reported issues in *ToysRUs* navigation having irrelevant things (e.g. gift cards). However, the page structure was poorly designed which required users to sequentially navigate through the page. Consequently, it is very difficult to address content problems without addressing the issues relating to the traversal of the website.

There is little information of how these problems influence blind users' experience on the web. Thus, it is difficult to suggest solutions that address these problems and actually improve users' experience. For instance, there were many problems with the poor page structure. It is unclear if the page structure problems are addressed whether participants will still have issues with the repeated effort due to the page refreshing. Also, it is not clear how much impact on users' performance has the lack of feedback if the page structure is addressed. A further investigation is required to see how specific design solutions can address these problems to improve users' experience on the web before design solutions are suggested.

Looking more closely at the problems encountered by blind users they map to the two design principles, feedforward and feedback (Norman, 1988, 2013). Problems with these two design principles can reduce the quality of users' experience. Feedforwards' role is to let users know what they can do on the website, how it works, what operations are possible, whereas feedback's role is to send information back to users regarding their actions, what happened. Norman (2013) notes that there are two parts to an action: execution and evaluation. When users face difficulties when they try to perform a task on a system/website they face two gulfs: the gulf of execution and the gulf of evaluation. Gulf of execution is when users encounter difficulties in figuring out how to operate a system, whereas gulf of evaluation is when users are having problems figuring out what happened after they performed an action. Norman (1988, 2013) argues that the feedforward and feedback design principles can help to bridge the gulfs of execution and evaluation, respectively. Problems with the poor page structure, interactive elements not being clear, headings not being descriptive or helpful and excessive effort required by users map directly to the feedforward design principle, which can result into the gulf of execution. The problem with the excessive effort required by users matches with Norman's explanation of gulf of execution "one measure of this gulf [execution] is how well the system allows the person to do the intended actions directly, without extra effort… (p. 51)" (Norman, 1988). The problems with the lack of feedback on users' actions map with the feedback design principle, which can result into the gulf of evaluation. Bridging the gulfs of execution and evaluation will make things visible for execution and evaluation.

Many of these problems were in relation to the filtering and browsing search of the page content. This is a very specific website design aspect, which, however, is present in a substantial number of websites. This design aspect is typically part of shopping category websites. Online shopping not only provide convenient service to sighted users but it has invaluable benefits to blind users. Previous research showed that using shopping websites can be a real challenge for blind users (Buzzi, Buzzi, Leporini, & Senette, 2010; Giraud et al., 2011).

Even for this specific design aspect, specific design solutions need to be tested to check how they can improve blind users' experience. There was a variety of different problems encountered around this design aspect for blind users, which makes it difficult to understand how specific design solution can benefit users' experience. For example, many interactive elements did not have a clear description. Users required to put an excessive effort due to the refreshing of the page each time a filtering option was selected. The lack of feedback on users' actions can confuse and disorient users that change even happened on the page. There were also many problems that were related to the poor structure of the page content, particularly the lack of headings on specific sections. All these different problems make it unclear how specific design solution to the problems blind users have can benefit users into improving their overall user experience. Further research is required to test how specific design solutions can benefit blind users' experience.

Although the study provided a further understanding of the problem differences between blind and sighted users on the web, it has some limitations.

The first limitation is that some of the blind participants (4 out of 12) that took part in the study were participants that also took part in the previous study (see Chapter 4) of the thesis. This was because the recruitment from a small participant pool is a very challenging task. However, there are a few reasons to believe that the participation in the previous study as well, did not have any impacts on the results of this study. First, there was a 10-month gap between the time the two studies took place. Also, the scope of the two studies slightly differ, with the present study focusing on the problems users identified, whereas in the previous study the experience of using the protocol was also investigated. To further support the assertion that the participation in the previous study did not impact the results, the mean number of problem instances by the participants that took part in the previous study (M = 31.25, SD = 17.86) and the participants that did not take part in the previous study (M = 32.00, SD = 14.67) were checked. As can be seen from the mean number of problem instances, the two user groups had a similar number of problems.

There are also some threats to the external validity of the results. The websites used in the present study can mainly be classified as shopping websites. Shopping websites are typically having a results section and a filtering section. Users have to interact with the filtering options of the page and then navigate through the page content to find the information they want. As in the previous study (see Chapter 4), users from both groups encountered many interactivity problems, it was preferred to use websites and tasks that would require users to interact with the various interactivity elements as well as navigate through the page content. Thus, the websites selected were all shopping websites and the tasks selected required both navigation between and within the pages but also data input. A possible limitation of the study is that some of the problems users encounter may be explicit to shopping websites, for example, the problems with the lack of feedback on users' actions when users select a filtering option. However, problems such as the lack of headings, links not being descriptive or images without appropriate alternative text are problems that can be found in other type of websites, such as library websites (Byerley & Beth Chambers, 2002).

Another threat to the external validity of the results is that the study included only desktop websites. The results of the study may not be generalised to mobile websites, with some problems that users had may be explicit to desktop websites. For example, blind users found issues with the skip links on the page. The skip link is often used to bypass blocks of content that are repeated in multiple pages (Caldwell et al., 2008). However, on mobile devices skip links are not often used. Also, it is not of the suggested accessibility guidelines of W3C that can apply to mobile devices (Patch, Spellman, & Wahlbin, 2015). Thus, some of the problems users had may not apply on mobile devices. It would have been interesting if future research would investigate the problem differences between blind and sighted users on mobile devices and check what of the problem differences identified in the present study apply on mobile devices as well.

The most important threat to the external validity of the results is the low number of websites used in the study. This study was only conducted with three websites. A larger number of websites was not preferred as each website evaluation for blind users required around an hour to complete. Thus, the number of websites selected was limited to three websites to mitigate fatigue effects. It is unclear whether the websites selected are representative of what users find on the web and whether the problems found related to the filtering options of the websites can be generalised to other websites. A further understanding of how common these features are on other websites, before investigating design solutions for this specific design feature, is required.

There were some limitations of the study that lowering its ecological validity. Participants did not use their own equipment. The rationale of participants not using their own equipment lies in the inability of assessing whether the equipment was going to be in running order before the arrival of the participant. Also, screen recording software that captures the session was preinstalled on the computers that were used in the present study. To mitigate any impacts of the participants not using their own equipment, they were asked to configure the equipment to their own preference in order to match their usual setup. Moreover, participants first performed a practice task, that was not analysed in the study, which made them more comfortable with the equipment they were using.

Regarding the measures collected, the study did not look into participants' efficiency or users' keystrokes. Regarding participants efficiency, no data were collected as it would not provide any new contribution to knowledge (Coyne & Nielsen, 2001; Disability Rights Commission, 2004). Regarding users' keystrokes, they could have been valuable data, if users' navigation techniques were investigated. However, for eliciting problems on the web keystrokes are rarely collected (Coyne & Nielsen, 2001; Disability Rights Commission, 2004; André Pimenta Freire, 2012; Lazar et al., 2012; Petrie & Kheir, 2007) and also they are not of much value for understanding users problems on the web (Dumas & Redish, 1999).

5.5. Conclusions

While there is an indication from the literature that the problems between blind and sighted users on the web differ, with a small overlap of problems between the two user groups, there is little detail in the literature of what these problems are and what causes them. Thus, the main goal of the present study was to investigate the problem differences between blind and sighted users on the web in order to provide a further understanding of what these problems are and what causes them. The study showed that there are problem categories distinct to each user group. For example, technology problems are distinct to blind users, whereas physical (visual) presentation problems are distinct to sighted users. Moreover, there are problem categories that have instances by both users, and these include the problems with the content on the page, problems with the information architecture of the page and problems with the users' interactivity with the page content. Most problem types of these categories had instances by both user groups. One of the main findings of this study is that the two user groups may encounter a similar type of problems, but the actual characteristics of the problems largely differ. There were also a few problem types that were distinct to blind users. This mainly included the problems related to the structure of the page and the lack of feedback on users' actions.

This result highlights a few important things. As the problems the two user groups encounter largely differ, this shows that addressing the problems sighted users have on the web would not necessarily mean that the problems blind users have will be addressed. This shows the importance of including blind users during the evaluation sessions of the websites. Also, it shows gaps in current website designs that do not accommodate the needs of blind users.

The work contributes to the existing knowledge of the problems that blind users encounter on the web by extending our understanding of how problems are distinct between blind and sighted. The study provides a further understanding of the range and diversity of user problems. It is important to understand the differences of the problems encountered by blind and sighted users on the web, in order to create solutions that properly address these issues. Given that blind users encounter a large number of problems that may be interacting with one another to impact users' experience, it is not clear which ones are having the impacts we are seeing on users' experience. Many of these problems were related to the search and filtering browsing of the page content. Although this design aspect is very specific, it is present in a considerable number of websites and it worth a further investigation for providing a design solution. However, before further looking into design solutions that address the problems around this specific design feature, there is an important study design limitation that needs to be considered. The study was only conducted with three websites. It is unclear whether the problems found around this design aspect can be generalised to other websites. Further investigation of the design features of a similar type of websites is required, to assess whether design solutions around this specific design feature are worth investigating. Thus, the next chapter of this thesis work presents an analytical study of how common are the design features of the websites used in the present study with a similar type of websites to assess the external validity of the results.

Chapter 6. Analysis of websites' features

6.1. Introduction

The results of the previous study demonstrated that there is a variety of problems that are impacting blind users. Many of the problems blind participants encountered are in relation to the search and filtering browsing of content, which is present in a substantial number of websites, particularly shopping websites. However, the study was conducted with only three websites, and each of those websites had search and browse features on them. It is possible that these websites are not representative of websites of this type, and thus the results of the previous study may not actually represent those types of problems and impacts the external validity of the study. Even though they are clearly impacting users, if they show up on few websites, then the benefits of further investigating design solution for these problems would be questionable.

Therefore, this study is set out to investigate whether or not a similar type of websites (i.e. shopping websites) have common design features with the websites used in the previous study. The analysis included a quantitative evaluation of websites' structure and interactive features that are related to the causes of the problems blind users found around the search and filtering browsing of content. The study addresses the following research question:

- Do the websites used in the study in Chapter 5 have common design features with websites of a similar type?

6.2. Method

The method employed was a content analysis which is consistent with previous research used to identify and quantify structural and functional features of websites (Herring, Scheidt, Wright, & Bonus, 2005). The approach followed comprised of three phases: preparation, organisation and reporting (Elo & Kyngäs, 2008).

The preparation phase involved the unit of analysis, which in this study was the page of the website that included filtering options and a list of results. A representative sample of websites was selected to perform the feature analysis. The sample was selected from Alexa web service, a website that offers statistics of websites' traffic and ranks the websites based on their popularity. The websites are ranked either globally, by country or by category. From the available website categories offered by Alexa, the one that closest match to the websites used in the previous study was the shopping category. For example, shopping websites include filtering options and results section, which are in line with the design features of the websites in the present study. Alexa provides a list of the top 500 shopping websites ranked based on their average daily visitors and page views over the last month. The inclusion criteria were that the websites used English and did not require any paid subscription to access their content. The sample of websites selected were the first 20 shopping websites as ranked by Alexa on April 12th, 2016.

Table 28 shows the sample of websites selected. In the 3rd and 4th rank was Netflix and Amazon (UK), respectively. However, both websites were excluded, as Netflix requires a paid subscription to access its content and Amazon's (UK) website was found to be identical with Amazon.com (which is already in the list) when it was checked regarding its design features.

Table 28. Top 20 shopping websites as ranked by Alexa on the 12 April 2016.

Shopping Website	**Alexa Ranking**
Amazon	1
Ebay	2
Walmart	5
Etsy	6
Target	7
Ikea	8
Homedepot	9
Steam	10
Groupon	11
Bestbuy	12
Lowe’s	13
Macy’s	14
H&M	15
Kohls	16
Nike	17
Newegg	18
Gap	19
Costco	20
Nordstrom	21
Wayfair	22

The organisation phase involved the creation of a coding scheme. The coding categories included two major categories: structure and interactivity.

The structure category included the use of headings to structure the page content and the order users perceive the content. These specific structure criteria were checked as they are related to structure problems reported in the previous study. For example, blind participants reported many issues with missing headings to indicate major sections of the page, such as the filtering section or the results section. Also, participants reported problems with the results in the list of results not marked as headings. Participants also reported issues with the headings of the page not being in hierarchical order, meaning lowest level headings were not contained within higher level headings. Another structure problem reported by participants was the order they encounter the page content. For example, participants were encountering the page content not in appropriate order.

The interactivity category included the approach used to update the page content. For example, whether the page updated its content as soon as users select a filtering option. Also, the feedback provided on users' actions was investigated. These two interactivity features were investigated as they relate to the interactivity problems blind users encounter in the previous study. For example, blind participants reported issues with the lack of submit button and the refreshing of the page each time a filtering option was selected. Also, participants reported issues with the lack of feedback on their actions, as they were unsure whether their actions were performed successfully.

To support the investigation of some of the features, a screen reader was used. The screen reader used was VoiceOver that comes with El Capitan OS, on Safari browser on a MacBook Pro. Also, the Web Developer toolbar plugin and the JavaScript Bookmarklets for Accessibility Testing on Chrome browser were used.

The last phase of the process involved an inter-coder reliability of a sample of four websites from another researcher with more than five years of experience in human-computer interaction methodologies and web design and development. Cohen's Kappa revealed a satisfactory level of agreement with K = 0.855 (Landis & Koch, 1977).

6.3. Results

This section presents the results of the comparison of the design features between the websites used in the previous study (Chapter 5) and the top 20 shopping websites as ranked by Alexa. Each design feature analysis is presented in a table that includes a column for the top 20 shopping websites and a column for the websites of the previous study, and presents the frequency of the different approaches used across the websites in each set.

6.3.1. Structure features

When considering the structure features, the study first considered an analysis of the use of headings on the websites. Table 29 lists the design approaches used for structuring the page content in the two sets of websites.

Table 29. Use of headings by the two sets of websites.

Design Feature	Top 20 websites N (%)	Websites from the previous study N (%)
Use of headings on the page	20 (100%)	3 (100%)
Headings are in a hierarchical order	2 (10%)	0 (0%)
Headings to indicate the filters section	4 (20%)	0 (0%)
Headings to indicate each filter option	9 (45%)	3 (100%)
Headings to indicate the results section	8 (40%)	1 (33.3%)
Headings to indicate each result	7 (35%)	1 (33.3%)

As presented in the table above, all websites in both sets used headings to structure the page content. From the 20 shopping websites, only two (10%) websites had the headings in hierarchical order, whereas from the websites of the previous study none of them had the headings in hierarchical order.

The analysis of headings showed that more than half of the websites in both sets of websites did not include headings to indicate major sections on their pages. For example, very few websites used headings to indicate the filtering section from the top 20 shopping websites, whereas none of the websites from the previous study used any heading. The number of websites indicating each filtering section is higher, with almost half of the websites from the top 20 shopping websites and all the websites from the previous study used headings to indicate each filtering option. For the results section and each result in the list, the figures are lower in both sets with less than half of the websites using headings to indicate these sections.

Another structure feature investigated (see Table 30) was the order that blind users perceive the page content.

Table 30. Order that users perceive the content by the two sets of websites.

Design Feature	**Top 20 websites N (%)**	**Websites from the previous study N (%)**
Filters – Results	16 (80%)	3 (100%)
Results – Filters	4 (20%)	-

The most common approach from both sets of websites was first presenting the filtering options and then the list of results. A further look into the order blind users perceive the content on the page showed that all websites that first had the filtering section and then the results section were matching with the visual page order. For the four websites from the top 20 shopping websites that first had the results and then the filtering options, it was found that three of them did not match the visual order of the page content and the order that blind users perceived the page content.

6.3.2. Interactivity features

For the interactivity features, an analysis of the design features of websites that were the causes of the key problems blind users found in relation to the search and filtering of page content was performed. This analysis included the extra effort required and the lack of feedback on users' actions.

Table 31 lists the design features of websites that are related to the extra effort.

Table 31. Design features for updating the page content by the two sets of websites.

Design Feature	**Top 20 websites N (%)**	**Websites from the previous study N (%)**
Update results using a submit button – page refresh	0 (0%)	0 (0%)
Update results using a submit button – focus in place	0 (0%)	0 (0%)
Update results on filter selection – page refresh	13 (65%)	3 (100%)
Update results on filter selection – focus in place	7 (35%)	0 (0%)

As is presented in Table 31, none of the websites provided a submit button to initiate a change of context. All websites, in both sets, were updating their page content as soon as users were selecting a filtering option.

Two-third of the websites from the top 20 shopping websites and all websites from the previous study updated their page content with the page refreshing. This design approach causes the screen reader focus being at the top of the page each time a change in context is requested. The other third of websites, from the top 20 shopping websites, update the page content with the screen reader focus staying in place.

The other interactivity feature investigated was the feedback on users' actions. The results of the analysis are presented in Table 32.

Table 32. Design features for providing feedback on users' actions by the two sets of websites.

Design Feature	**Top 20 websites N (%)**	**Websites from the previous study N (%)**
Update of the page title with the filtering options – page refresh	3 (15%)	0 (0%)
Update of the page title with the filtering options – focus in place	0 (0%)	0 (0%)
Immediate feedback on users' actions when a change of context requested – focus in place	0 (0%)	0 (0%)
Summary of selected filters provided	14 (70%)	2 (66.6%)

Almost all websites in both sets lack an immediate feedback on users' actions. The only feedback provided on many websites was a summary section of the selected filtering options. However, very few websites had a heading to indicate this section. For the top 20 shopping websites only two (14%) out of the 14 websites, whereas none of the websites from the previous study used any heading to indicate this section.

The only form of instant feedback on users' actions found was the update of the page title to include the selected filtering options each time a change of context requested. This approach was consistently found on three websites from the top 20 shopping websites. Another two websites updated their page title, but it was not consistent across all filtering options. For example, Macy's did not update the page title when the price filter was selected, whereas Nike did not update the page title when the size filter was selected. None of the websites from the previous study updated their page title when users' request a change of context.

For the websites that updated the page content and the screen reader focus stayed in place, there was not any form of instant feedback for blind screen reader users to inform them that an update of the page content has happened.

6.4. Discussion

An analysis of the features of a similar type of websites was conducted to investigate the extent to which the problems found by blind users in relation to the search and filtering browsing of content in the previous study (Chapter 5) can be generalised to other websites. The websites selected for the analysis were all classified as shopping websites by Alexa. The top 20 shopping websites were selected as ranked by their popularity, meaning the websites with the most visitors were selected. The analysis involved an investigation of common structure and interactivity design features between the two set of websites. The features investigated included specific design aspects that were the causes of the main problems blind users encountered in the previous study in relation to the search and filtering browsing of websites' content.

The analysis showed that there were many similarities between the two set of websites for the structure design features. All websites, in both sets, used headings to structure the page content. However, very few websites in both sets had headings to indicate specific sections of the page. The lack of headings can cause difficulties to blind users navigating the filtering and results section of the page effectively as it was found in the previous study.

A common structural design feature was the use of headings not in a hierarchical order. The use of headings in a hierarchical order is quite important as blind users often use the headings to navigate through the page content and get an overview of the overall structure of the page content (Power et al., 2013; WebAIM, 2014). If the headings are not in a logical order blind users may have difficulties to navigate and find particular information.

There were also similarities between the two sets of websites for the order that users perceive the content on the page. Most websites in both sets first present the filtering options and then the results list. Moreover, they match the visual presentation of the page content with the order blind users perceive the content, meaning blind users will experience the page content in the correct reading order as sighted users.

The two sets of websites were very common in regard to their structural design features. That means many of the problems blind users encounter in relation to the page structure, particularly the lack of headings to structure and organise the page content, can be generalised to other shopping websites.

The analysis also showed similarities for the interactive design features between the two sets of websites. First, none of the websites in both sets provided a submit button to initiate a change of context, as the pages were updating on users filtering option selection. This design approach seems to be the norm for search and filtering browsing websites. Another common feature found was that most websites update their page content by refreshing the page, which takes the screen reader focus at the top of the page each time a change of context is requested. This design feature was the cause of blind users' excessive effort problem that was extensively reported in the previous study. Also, there were some websites that update the page content with the screen reader focus staying in place. This approach may be considered more appropriate for blind users, as addresses the problems with the extra effort required due to the refreshing of the page. However, there is an important limitation that needs to be considered. Blind users will not know that their action was performed on the page. None of the websites that used this approach provided any form of instant feedback on users' action that the page content updated.

Another similar feature was the lack of instant feedback about users' actions in both sets of websites. There were a few websites from the top 20 shopping websites that provided feedback on the title of the page. However, there are a few drawbacks of this approach. First, it is not clear whether this approach will work with blind screen reader users as it has not empirically tested if it benefits users' experience. Second, it will only work when the page refreshes on users' actions, as the screen reader only reads the page title when the page refreshes. Third, users may not perceive the feedback added to the page title as it not a standard design approach. Also, the position that the feedback added can influence its identification. To demonstrate this approach with an example, Amazon updates the page title with the selected filtering options added at the end of the page title. When the page refreshes, the screen reader will start reading the page title. However, users must stay inactive for the screen reader to read all the page title content. If users press a button to navigate into the page content, their reading of the page title will be interrupted. Another limitation of this approach is that the page title in some websites included only the filtering option selected rather than including the name of the filter as well. For example, in the Newegg when users select the 500GB hard drive, the page title adds “500 GB” on the existing page title. However, the page offers options for both hard drive disks and solid-state drives. It is not clear which of the two filters the selected option is referring to. Moreover, the purpose of the page title is to help all users to quickly and easily identify whether the information contained on the page is relevant to their needs. Providing feedback on users’ action on the page title can produce too long page titles that may cause difficulties to other user groups. For example, sighted users can only see the part of the title that is displayed on the tab panel of their browser. If the page title is starting with the filtering option, they may have difficulties in identifying the correct page tab when they have multiple pages open on their browser.

Another common feature was the provision of a summary of the selected filtering options on the page. However, most websites in both sets did not provide a heading to indicate the summary section, which can make it difficult to find it on the page.

The results of the analysis showed that there are many common design features between the two set of websites. This can imply that the problems blind users encountered in relation to the search and filtering browsing of content in Chapter 5 can be generalised to other shopping websites. Shopping websites are of particular interest for blind users, as they may not be able to shop on their own at the stores. However, shopping online can be a challenging activity for blind users. A survey of users online shopping experience showed that blind users shop much less than sighted users online due to the difficulties they encounter. Blind users reported that have problems navigating and receiving adequate information when searching and choosing products (Buzzi, Buzzi, Leporini, et al., 2010). This is further supported by other empirical studies with blind users encountering many problems in shopping websites (Giraud et al., 2011; Stenitzer et al., 2008). Based on the results of the analysis, many of the problems blind users had in the previous study can be generalised to other shopping websites and previous research that points to the difficulties blind users have in shopping websites, they motivate further research on testing different design solutions that can help to improve blind users' experience in search and browse websites.

6.5. Conclusions

This study set out to determine whether the problems found by blind users in relation to the search and filtering browsing of content in the previous study can be generalised to other shopping websites.

To investigate this notion, an analysis of the common structure and interactivity design features of the websites from Chapter 5 and the top 20 shopping websites as ranked by Alexa was conducted.

The study showed many design similarities between the two set of websites, which implies that many of the problems reported in the previous study can be generalised to other shopping websites. All websites had issues with their structure and lacked instant feedback on users' actions. Moreover, the most common approach for updating the page content had the page refreshing, which causes the screen reader focus going at the top of the page. This approach is related to the excessive effort problem reported by blind users in the previous study.

The problems found by blind users' around this specific website design feature seems to influence each other, which makes it difficult to understand the benefits of design solutions to these problems without empirically testing them. To get a better understanding of how specific design solutions can address these problems to improve the effectiveness and efficiency of users and the perceived usability of a website further research is required. The next chapter of this thesis proposes different design solutions to some of the key problems blind users had and how to evaluate them in order to explore how and if they benefit blind users' experience.

Chapter 7. Proposing design solutions to the key problems of blind users

7.1. Introduction

The results of the study in Chapter 5 demonstrated that there is a variety of problems that are impacting blind users. Many of the problems blind participants encountered are in relation to the search and filtering browsing of content, which is present in a substantial number of websites. These problems seem to present on many shopping websites (see Chapter 6).

To improve blind users' experience in search and filtering browsing of content websites a set of three website designs have been proposed and implemented. This chapter presents the proposed design solutions and the experience between each website design for blind users.

7.2. Proposed design solutions

Looking at the search and filtering browsing of content problems, there could be a number of different causes of the problems. The problems discovered in Chapter 5, could relate to the fact that the websites used had many structural issues. The presence of problems relating to information architecture on the page may prevent users from finding and interacting with the filtering options entirely. However, if these issues are solved, it is unclear if users would be more successful. A number of problems reported were related to users having to put in excessive effort to retrace their steps within pages when the page refreshed due to the requested change of context. Other problems involved users having problems assessing the state of the website when a changed happened on the page, as they were not getting any feedback on their actions.

On review of the problem types reported, there were three key problems that occurred frequently that are independent of the information design and content of the website and potentially contribute to users' lack of success in their tasks:

- A lack of structure, which prevents users from navigating the filtering and results sections of the page effectively.
- A lack of feedback relating to what was happening on the page in relation users' actions.
- An excessive effort required by users' due to the refreshing of the page each time a filtering option was selected. Users were being at the top of the page and had to retrace their steps to where they were interacting with the screen reader.

These problems seem to influence each other. A poor page structure will make it difficult for users retrace their steps to where they were interacting with their screen reader when the page refresh, which can possibly augment the perceived extra effort. A complete lack of feedback will leave users disoriented that change even happened. As a result, it is difficult to understand how these problems influence the effectiveness and efficiency of users and the perceived usability of a website.

Based on these problems, a set of three website designs were proposed with progressively improving the website design. All three website designs were implemented using the same web technologies.

7.2.1. General Website Design

Each website design was implemented using HTML5, CSS3, JavaScript and the Accessible Rich Internet Applications (ARIA) specifications (Diggs, McCarron, Cooper, Schwerdtfeger, & Craig, 2017). Besides, PHP and MySQL were used to build the search and filtering browsing of the content functionality of the websites. In addition, the proposed techniques by WCAG 2.0 success criteria (Caldwell et al., 2008) that were applicable to the system were followed. Each website was created to have identical content structure. Specifically, it had four main areas: the header area, the filtering section, the results section and the footer. Each of the websites was designed so as to actual data could replace the filtering options and the content, allowing for users to use each website as a new website to avoid any familiarity effects.

7.2.2. Structure only design intervention

The first design intervention is addressing the problems in relation to the structure of the page. It puts landmarks, all content areas are well labelled with appropriate headings, in the page to support screen reader users browsing strategies (Power et al., 2013; T. Watanabe, 2009; WebAIM, 2014). Figure 27 shows the document outline, the structure of the website generated by the headings of the website. As can be seen from the figure, all major sections of the page are indicated using headings. The filters, selected filters and results sections are indicated using a level 2 heading, whereas the child elements of each of these sections are indicated with level 3 headings. For example, each filtering option and each result in the list of results is indicated using a heading level 3. Having headings to each content areas is addressing one of the main problems found by blind users in the study in Chapter 5. Also, the headings are in a hierarchical order without any skipping levels. That means participants will not be navigating from a level 2 to level 4 heading. Both aforementioned problems with the website headings were also found in other shopping websites (see Chapter 6). Moreover, the order of content that blind users were perceiving matches the visual presentation of the page content, meaning blind users will experience the page content in the same reading order as sighted users. Blind users were first finding the filtering section and then the results section, a common approach that was found in the websites used in Chapter 5 and the websites analysed in Chapter 6.

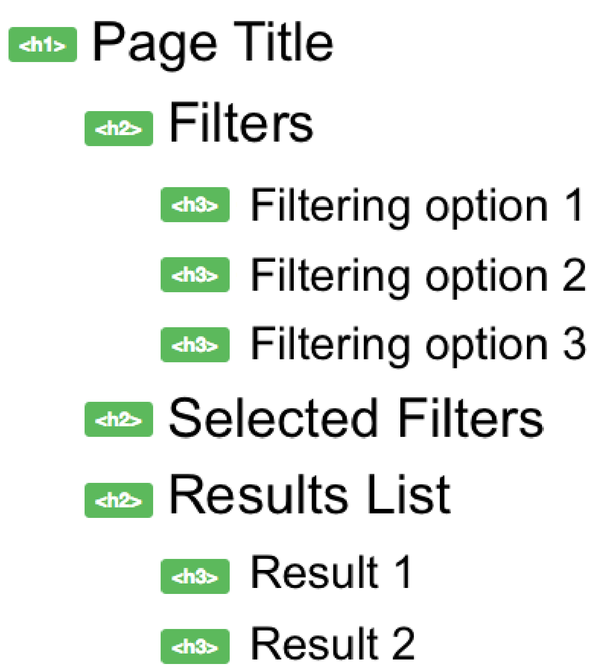

Figure 27. Document Outline generated by the headings of the page using the Web Developer plugin on Chrome browser.

All form controls had associated labels and all links were well labelled with their purpose being clear from the link itself, a technique that was found to work better for blind users for identifying the destination of the link (Power et al., 2011). Also, a summary of the selected filtering options was available on the page indicated by a heading.

The page was updating its content each time uses selected a filtering option, without providing a submit button. The rationale of this approach lies in the fact that it seems to be the norm on this type of website. This approach was used on all three websites in the study in Chapter 5 and all websites from the analysis in Chapter 6.

The following transcript of a screen reader output shows the experience of using the website design. In this example, users need to find the number of cities that are in Europe. The transcript below shows the experience of navigating through the page content via a screen reader (VoiceOver) in linear order. As can be seen, when users select the filtering option Europe (transcript line 8), the page refreshed. Users' screen reader focus goes to the top of the page. Then, they have to navigate again through the page content to assess the status of the page.

Transcript showing the experience navigating through the structure only website design

SR: screen reader output, UA: user's action

1. SR: Wopolis HTML Content
2. SR: heading level 1, link, Wopolis
3. SR: heading level 2, Filters
4. SR: heading level 3, Continent

5. SR: Asia, unchecked, checkbox
6. SR: Asia
7. SR: Europe, unchecked, checkbox
8. UA: Selection of Europe
9. SR: Sound that the page refreshed
10. SR: heading level 1, link, Wopolis
11. SR: heading level 2, Filters
12. SR: heading level 3: Continent
13. SR: Asia, unchecked, checkbox
14. SR: Asia
15. SR: Europe, checked, checkbox
16. SR: Europe
17. SR: North America, unchecked, checkbox
18. SR: North America
19. SR: Africa, unchecked, checkbox
20. SR: Africa

…

21. SR: heading level 3, Official Language
22. SR: Arabic, unchecked, checkbox
23. SR: Arabic
24. SR: Chinese, unchecked, checkbox
25. SR: Chinese

…

26. SR: heading level 3, Population
27. SR: less than 100,000, unchecked, checkbox
28. SR: less than 100,000

...

29. SR: heading level 2, Selected Filters

30. SR: Continent: Europe

…

31. SR: heading level 2, Cities List

32. SR: Number of Cities: 841

This design approach, addressing the problems in relation to the structure of the page, was adopted by all website designs implemented. Given that previous work (Power et al., 2013; T. Watanabe, 2009) showed that the addition of headings within the page provides obvious benefits to blind users, implementing an unstructured page as a control condition was not seen to provide any value in terms of a new contribution. The structure of page content should provide screen reader users with an easy and quick navigation through the page content (T. Watanabe, 2009) using a standard navigation strategy (Power et al., 2013). This design, for all intents and purposes, is the most common design approach with a structured content which is expected a competent accessibility knowledgeable designer to be able to produce.

7.2.3. Structure and feedback at the top of the page design intervention

The second design builds on the website design structure only by further incorporating feedback that informs users of what happened in response to their actions at the top of the page. Similarly, with the structure only website design, the page updates its page content as soon as users select a filtering option. The only difference is the extra feedback provided at the top of the page that informs users what was the result of their action.

Figure 28 shows an example of what it was like to experience the feedback using a screen reader. In the example below users need to find the number of cities in Europe continent.

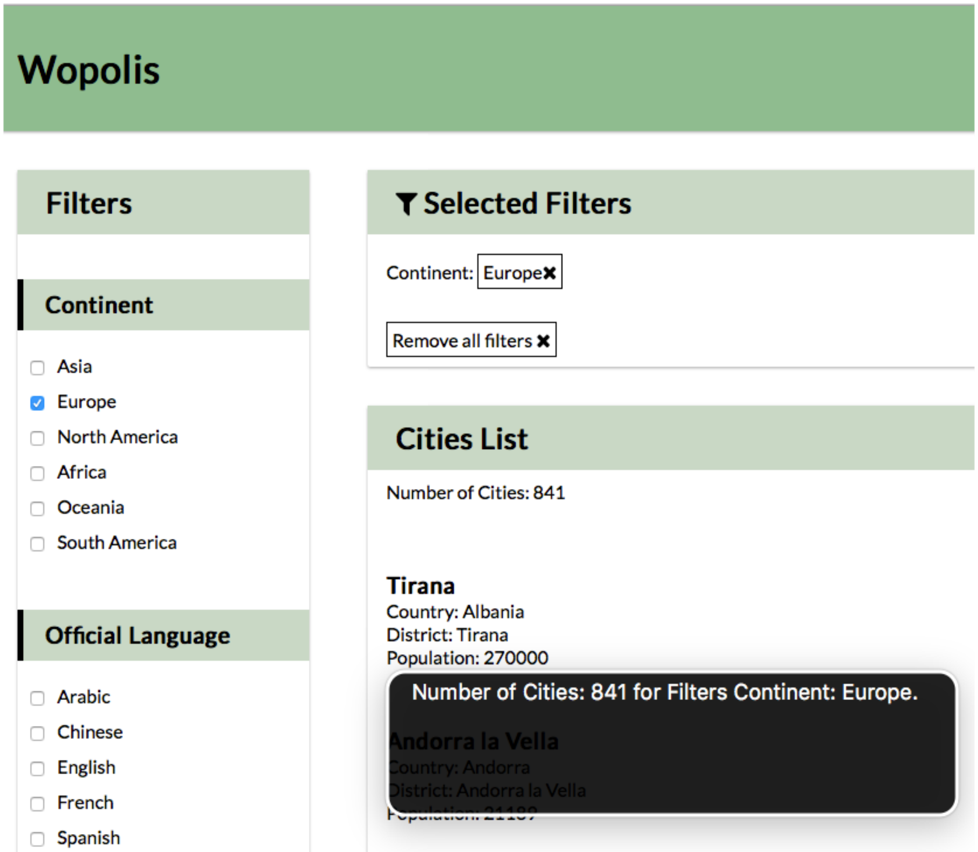

Figure 28. Example of how the instant feedback is provided in structure and feedback at the top of the page website design.

In this website design, when a user requests a change in context by selecting a filtering option, the page will refresh with the focus of the screen reader going at the feedback regarding the current state of the page. This will be the first information encountered on the page by the screen reader when the page refreshed. This is non-standard design, that is currently not commonly implemented, but it provides a minimal amount of instant feedback that is easily available to users without having to navigate through the page content. If the feedback regarding the outcome of users' actions is useful, it may reflect with improvements on users' experience on this website design.

The following transcript of a screen reader output shows the experience of using this website design. Using the same task as the transcript in the previous website design, the user needs to find the number of cities that are in Europe. Thus, the user needs to select Europe in the Continent filter options. The transcript below shows the experience of navigating through the page content via a screen reader in linear order. As can be seen, when users select a filtering option, Europe, the page refreshed (transcript line 9). The user goes to the first element of the page which is the instant feedback regarding the state of the page. In this case, it was: "Number of Cities: 841 for filters Continent: Europe." (transcript line 10). This instant feedback on users' actions it is believed to make it clearer what is happening in the page in response to users' actions.

Transcript showing the experience navigating through the structure and feedback at the top of the page website design

SR: screen reader output, UA: user's action

1. SR: Wopolis HTML Content
2. SR: heading level 1, link, Wopolis
3. SR: heading level 2, Filters
4. SR: heading level 3, Continent
5. SR: Asia, unchecked, checkbox
6. SR: Asia

7. SR: Europe, unchecked, checkbox
8. UA: Selection of Europe
9. SR: Sound that the page refreshed
10. SR: Number of cities 841 for filters Continent: Europe
11. SR: heading level 1, link, Wopolis
12. SR: heading level 2, Filters

This website design intervention does not require any visual change on the website, as the feedback message is visually hidden. Thus, there is not any visual difference with the structure only design as the extra feedback provided is only perceivable via a screen reader. This design intervention will not affect the visual design of websites anyhow if they adopt it.

This approach was preferred rather than using the title of the page to provide feedback on users' actions. In the analysis of the top 20 shopping websites (see Chapter 6) it was found that a few websites used the title of the page to provide feedback on the selected filtering options. However, this approach is not empirically tested if it can benefit blind screen reader users' experience. Moreover, users may not perceive the extra feedback provided as it is not a standard design approach. In addition, the websites that provided feedback on the title of the page only included the filtering options selected. However, this approach can cause other problems as there may be instances where filtering options in different sections may have the same name, as it was demonstrated in Chapter 6. Moreover, the title of the page should be used to help users quickly and easily identify whether the information contained on the page is relevant to their needs. Using the page title to provide feedback on users' actions can produce too long page titles that may be difficult to be distinguished by sighted users when they are opened on their browser tab panel. Thus, it was preferred to provide the additional feedback at the top of the page rather than using the title of the page.

7.2.4. Structure and in-place feedback design intervention

The third website design intervention was also implemented using the structure only as a template. However, this design updates its content on users' request without refreshing the page, which will allow users to continue their tasks from where they were left off without any extra effort.

This design approach was also found by a few websites in the study in Chapter 6. However, a problem with the approach used in those websites was that blind users do not get any feedback that a change of context happened. To address this issue, in this website design the screen reader will provide in-place feedback regarding the filtering options selected and the number of results returned. This was achieved using WAI-ARIA live regions (Diggs et al., 2017), which allows screen readers to perceive areas that have been updated, in this case, the state of the page.

Figure 29 presents an example of what it was like to experience the feedback using a screen reader in structure and feedback in-place website design. In the example below users need to find the number of cities in Europe continent. When the users select Europe from the continent options, the page updates its content and provides feedback on users' actions. However, the screen reader focus stays in place (at the selected filtering option).

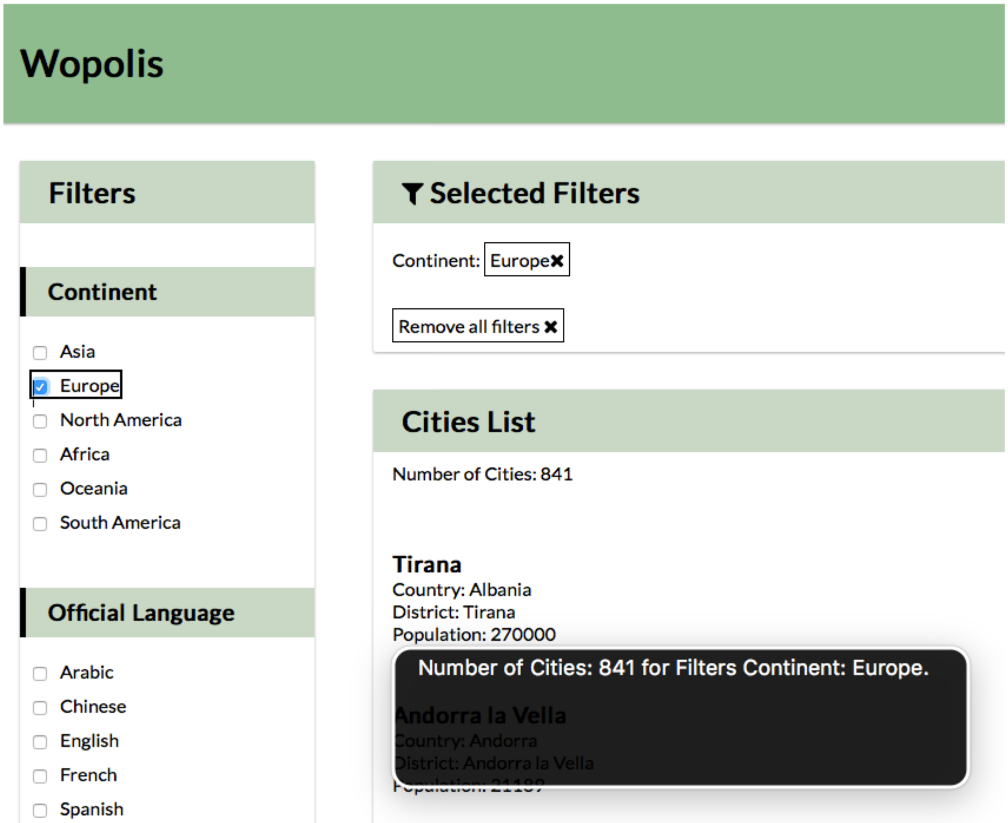

Figure 29. Example of how the instant feedback is provided in structure and in-place feedback website design.

To further understand the experience using this website design a transcript of the screen reader output is presented below. The transcript presents the experience of using the screen reader (VoiceOver) when users try to find the number of cities in Europe. As it can be seen, when users select the Europe checkbox (transcript line 10) the screen reader provides instant feedback on users' actions: "Number of Cities: 841 for Continent: Europe". Moreover, the screen reader focus stays in place (transcript line 11).

This website design is believed to address the problems with the lack of feedback on users' action but also the problems with the excessive effort required due to the refreshing of the page.

Transcript showing the experience navigating through the structure and in-place feedback website design

SR: screen reader output, UA: user's action

1. SR: Wopolis HTML Content
2. SR: heading level 1, link, Wopolis
3. SR: heading level 2, Filters
4. SR: heading level 3, Continent
5. SR: Asia, unchecked, checkbox
6. SR: Asia
7. SR: Europe, unchecked, checkbox
8. UA: Selection of Europe
9. SR: check, Europe, checkbox
10. SR: Number of Cities: 841 for filters Continent: Europe
11. SR: Europe
12. SR: North America, unchecked, checkbox
13. SR: North America

This is the more complex design solution as it combines several design solutions, good structure, feedback on users' actions and does not refresh the page on users' actions. Also, this website design intervention does not require any visual change on the website, as the feedback message is visually hidden. Thus, there are not any visual differences with the other two website designs.

7.2.5. Compliance of design interventions to accessibility guidelines (WCAG 2.0)

All websites were designed following the WCAG 2.0 AA success criteria (Caldwell et al., 2008). However, only the success criteria that were applicable for blind users were followed. For example, all content areas were labelled using appropriate headings, form controls had associated labels, all links were well labelled with their purpose being clear from the link itself, there was proper reading order.

All three website designs comply (see Table 33) to the same WCAG 2.0 success criteria. There were though a few success criteria that do not apply to the websites created as they relate to multimedia or visual requirements, such the colour contrast. The only success criterion that applies to the website designs but they do not comply to is 3.2.2 On Input Level A. The website designs fail this success criterion as they cause a change of context when an interface component change its setting. In all three websites designs when users select a filtering option, the page updates its content without providing a submit button for initiating a change of context. However, this approach seems to be the norm on these type of websites (see Chapter 6). For this reason, a submit button was not provided as it would not seem to be beneficial to test an approach that is not in line with common practices in search and browse websites.

Table 33. Conformance level of WCAG 2.0 AA of the three website designs.

Success Criterion	**Structure only**	**Structure and feedback at the top of the page**	**Structure and in-place feedback**
1.1.1 Non-text Content Level A	Pass	Pass	Pass
1.2.1 Audio-only and Video-only (Prerecorded) Level A	N/A	N/A	N/A
1.2.2 Captions (Prerecorded) Level A	N/A	N/A	N/A
1.2.3 Audio Description or Media Alternative (Prerecorded) Level A	Pass	Pass	Pass
1.2.4 Captions (Live) Level AA	N/A	N/A	N/A
1.2.5 Audio Description (Prerecorded) Level AA	N/A	N/A	N/A
1.3.1 Info and Relationships Level A	Pass	Pass	Pass
1.3.2 Meaningful Sequence Level A	Pass	Pass	Pass
1.3.3 Sensory Characteristics Level A	N/A	N/A	N/A
1.4.1 Use of Color Level A	Pass	Pass	Pass
1.4.2 Audio Control Level A	N/A	N/A	N/A
1.4.3 Contrast (Minimum) Level AA	N/A	N/A	N/A
1.4.4 Resize text Level AA	N/A	N/A	N/A
1.4.5 Images of Text Level AA	Pass	Pass	Pass
2.1.1 Keyboard Level A	Pass	Pass	Pass
2.1.2 No Keyboard Trap Level A	Pass	Pass	Pass
2.2.1 Timing Adjustable Level A	N/A	N/A	N/A
2.2.2 Pause, Stop, Hide Level A	N/A	N/A	N/A
2.3.1 Three Flashes or Below Threshold Level A	N/A	N/A	N/A

2.4.1 Bypass Blocks Level A	Pass	Pass	Pass
2.4.2 Page Titled Level A	Pass	Pass	Pass
2.4.3 Focus Order Level A	Pass	Pass	Pass
2.4.4 Link Purpose (In Context) Level A	Pass	Pass	Pass
2.4.5 Multiple Ways Level AA	N/A	N/A	N/A
2.4.6 Headings and Labels Level AA	Pass	Pass	Pass
2.4.7 Focus Visible Level AA	N/A	N/A	N/A
3.1.1 Language of Page Level A	Pass	Pass	Pass
3.1.2 Language of Parts Level AA	Pass	Pass	Pass
3.2.1 On Focus Level A	Pass	Pass	Pass
3.2.2 On Input Level A	Fail	Fail	Fail
3.2.3 Consistent Navigation Level AA	Pass	Pass	Pass
3.2.4 Consistent Identification Level AA	Pass	Pass	Pass
3.3.1 Error Identification Level A	N/A	N/A	N/A
3.3.2 Labels or Instructions Level A	Pass	Pass	Pass
3.3.3 Error Suggestion Level AA	N/A	N/A	N/A
3.3.4 Error Prevention (Legal, Financial, Data) Level AA	N/A	N/A	N/A
4.1.1 Parsing Level A	Pass	Pass	Pass

7.3. Evaluation of design solutions

The website designs will be evaluated in terms of the effectiveness (task success rates), the efficiency of users (task time), the perceived usability and the perceived workload.

The three website designs will be evaluated only by blind users. The rationale of this decision was that the three website designs do not have any visual differences. The instant feedback provided on both structure only and feedback at the top of the page and structure and in-place feedback website designs is only perceivable via a screen reader. In addition, the elimination of the extra effort required due to the refreshing of the page it is not believed to have any influence on sighted users' experience as the mouse cursor of sighted users will stay at the same position each time the page refresh. For these reasons, the evaluation of the website designs proposed will be conducted only with blind users.

The next chapter of this thesis presents an evaluation of the three website designs proposed by blind users.

Chapter 8. Empirical study of the benefits of specific design solutions on blind users' experience in search and browse websites

8.1. Introduction

The study presented in this chapter investigates how specific design solutions through progressively improving the website design can address common interactivity problems to improve blind users' experience.

A set of three website designs with progressively improved designs were implemented. The first design addresses the page structure, by creating a filtering browsing website that follows known accessibility guidelines (Caldwell et al., 2008). It puts landmarks, all content areas were labelled with appropriate headings, in the page to support screen reader users browsing strategies (Power et al., 2013; T. Watanabe, 2009; WebAIM, 2014). The second design incorporates instant feedback that informs users of what happened in response to their actions on page refresh. The third design takes this feedback and further improves on it by eliminating the page refreshing. It leaves the screen reader focus on users' position where they triggered the action.

By having the users completing tasks on these three different website designs and comparing a variety of different user experience measures, it will reveal the benefits of each design solution on users' experience. Also, it will check the differences of providing specific design solutions that solve the problems (i.e. providing only good page structure) and providing a more complex design solution that combines several design solutions.

The unified definition of web accessibility proposed in Chapter 3 was used to devise the study. In relation to the key components of the unified definition, the study manipulates the design of the website only. The measures collected included users' effectiveness, efficiency and satisfaction on each website design.

The study addresses the following research question:

- What are the benefits of specific design solutions to the key problems in search and browse websites on blind users' experience?

8.2. Method

8.2.1. Study design

A repeated measures within-participants design was used, with the website design as the independent variable with three levels (*structure only*, *structure and feedback at the top of the page*, *structure and in-place feedback*). The dependent variables were qualities of usability, such as participants task success rate, task time, ease of use, perceived workload and design preference.

8.2.2. Participants

Eighteen blind screen reader participants took part in the study. Fifteen of the participants were men and three were women. Ages ranged from 25 to 72 (M = 51.7, SD = 16.2). Nine of the participants were congenitally blind while the remaining nine lost their sight between the ages of 3 and 35. Due to the limited participants pool, 5 out of the 18 participants, were participants that took part in the previous study of this dissertation, of which one took part in the study in Chapter 4 as well.

None of the remaining 13 participants took part in any of the previous studies of this dissertation.

Participants rated their experience and expertise on the web using a five-point Likert item, where 1 means “very low” and 5 means “very good”. The average rating of web experience was 4.2 (*SD* = 0.7), whereas the average rating of web expertise was 3.8 (*SD* = 0.8).

All participants used screen readers to access computers and the web for home and work. Thirteen participants used JAWS (running on Windows OS), three used NVDA (running on Windows OS) and two used VoiceOver (running on Mac OS). The JAWS version used varied from JAWS 15.0 to JAWS 17.0 (the latter being the latest version of JAWS when the study was conducted). Participants who used NVDA used the latest version 2016.2.1. Participants who used VoiceOver used the latest version that comes with Mac OS El Capitan (the latest version of Mac OS when the study was conducted). Participants were asked to rate their experience and expertise using screen readers on a five-point Likert item, where 1 means “very low” and 5 means “very good”. The average rating for experience and expertise using screen readers was 4.4 (*SD* = 0.6) and 3.9 (*SD* = 0.8), respectively.

Sixteen participants used Windows and two participants used Mac OS. The majority of the participants who used Windows reported Internet Explorer as their primary browser, and all the participants who used Mac OS reported Safari as their primary browser.

8.2.3. Equipment and Material

For participants who used the Windows OS, the study was conducted using a desktop computer running Windows 10 with speakers and a keyboard. For participants who used the Mac OS, the study was conducted using a MacBook Pro running the El Capitan OS with speakers. In addition, blind participants were able to choose the screen reader software they were most familiar with, for example, JAWS, NVDA or to use the VoiceOver version that comes with El Capitan OS on Mac. The screen reader software that participants used was already declared during the recruitment process, and all installation of the software was already arranged properly before the arrival of the participants to match their home or work environment.

Participants did not use their own equipment as I wanted to ensure that all equipment was in running order before the arrival of the participant. Also, the sessions were recorded, using Morae 3.1 on Windows and ScreenFlow 4.0.3 on Mac OS, which were preinstalled on the computers used in the study. These recordings included audio and screen activity.

When participants completed all tasks on each website they were asked to complete a questionnaire about the website using a 5-point Likert items:

- Q1: How easy or difficult did you find the website to use?
- Q2: How confident or not confident are you that you completed the tasks successfully?
- Q3: How clear or not clear was it to you, what was happening in the page in response to your actions?

Then, participants were asked about how was it to use each website design. Participants were also asked to complete the NASA TLX, a subjective workload assessment questionnaire (Hart & Staveland, 1988). Finally, at the end of the session participants were asked to rank the websites in terms of which one they preferred to use more.

8.2.4. Website Designs

The three website designs were structure only, structure and feedback at the top of the page and structure and feedback in-place (see Chapter 7). Figure 30 shows an example of one of the website designs. Each of the websites was designed so as to actual data could replace the filtering options and the content, allowing for users to use each website as a new website to avoid any familiarity effects. The content was downloaded from the free databases available on the MySQL website.

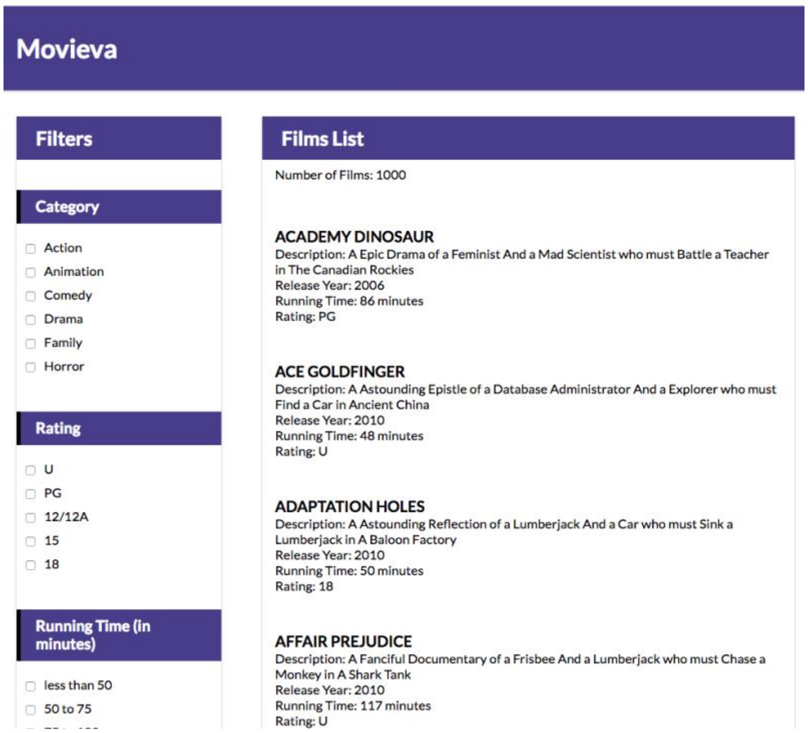

Figure 30. Example of a website design.

Three different websites resulted from this content change:

- Movieva[1] - a movie site
- GlobeTech[2] – a company site
- Wopolis[3] – a travel information site

8.2.5. Tasks

On each website, participants were given an introductory scenario and then they were asked to perform three tasks. Each of these tasks is designed to be know-item searchers, where the user attempts to find a specific piece of information that is guaranteed to be on the webpage. By eliminating the need to navigate between several different pages, as well as any content unrelated to the task within the website, these know-item searchers remove much of the irrelevant content users may encounter in an exploratory search.

Movieva

Scenario: You decided to go watch a movie with your partner at Movieva, a new cinema that opened last week in your town. However, you want to check the collection of movies they offer first. You decide to visit their website and check what movies they offer.

Tasks:

- What is the title and the description of the first listed film that is a Comedy and has a rating of 12/12A?

[1] Accessed from https://dev.mysql.com/doc/sakila/en/

[2] Accessed from https://dev.mysql.com/doc/employee/en/

[3] Accessed from https://dev.mysql.com/doc/world-setup/en/

- What is the title and the rating of the first listed movie that is a Family movie, is rated U, PG, or 12/12A and has a running time between 100 and 125 minutes?
- What is the title and the running time of the first listed movie that is an Action movie or a Horror movie, is rated 15 or 18 and has a running time of more than 150 minutes?

GlobeTech

Scenario: You are writing an article in your blog about the job prospects at GlobeTech, a technological company. Jenny suggested that you use their employees' list, which is available on their website in order to get more information about the employees of the company. You decided to visit the website to get information about their employees.
Tasks:

- What is the name and the gender of the first listed employee who is a member of Staff and is from the marketing department?
- What is the name and hire date of the first listed employee who is an Engineer, is in the Development, Production or Research departments and is paid between £40000 and £60000?
- What is the name and birthday of the first listed employee who is a Senior Engineer or is Senior Staff, is in any of the Development or Sales departments and earns more than £90000?

Wopolis

Scenario: You are planning your honeymoon for the next month and you want to go for holidays in another country. However, you want to learn more details about cities. Samantha suggested that you use Wopolis, an

online repository that contains information about all the cities of the world. You decided to visit Wopolis to get information about cities.
Tasks:

- What is the name and the population of the first listed city that is in Africa and has French as an official language?
- What is the name and the district of the first listed city that is in Asia, has English, Arabic or Chinese as its official language and has a population of 200 to 300 thousand people?
- What is the name and the country of the first listed city that is in Europe or Oceania, has English or Spanish as an official language and has a population of more than 500 thousand?

In preparation of the evaluation, the tasks were first undertaken using JAWS and NVDA on Windows and VoiceOver on Mac OS, to check that it is possible for screen reader users to complete the tasks.

8.2.6. Procedure

The study took place in the Interaction Laboratory at the Department of Computer Science of the University of York. Participants were briefed about the study and were asked to sign an informed consent form. To avoid any conflicts between the technology and participants' preferences, participants were asked which screen reader and browser they would like to use. Then, they were given the option to adjust the sound and related software to their preference in order to match their usual setup.

Participants first performed three tasks in a training website, which was not used in the analysis of the data. During the pilot study with a blind participant, it was noticed that the participant had a substantial training effect related to the structure of the website. Once the page structure was understood in one task, the others were substantially faster as all websites had the same structure. This had the potential to impact the results of any task after the first.

In order to minimise this learning effect, all participants performed three tasks in a separate pilot website which was a conformant structured page before they performed the tasks on the three websites of which data were collected for analysis.

The order that participants evaluated each website was counterbalanced using 3x3 (design x content) Latin Square. In total, nine websites were created to cover each combination of design and content. The tasks' order on each website was not counterbalanced, as the order of the tasks was of increasing difficulty based on the number of the steps required.

After participants had completed all the tasks in all the websites, they were asked to complete a questionnaire about their preference. Afterwards, participants were asked to complete a demographic questionnaire. Then participants were debriefed about the study, and I answered their questions. Finally, any information that was necessary for the compensation of participants' time was collected.

8.2.7. Data Analysis

The video recordings of each participant were reviewed, in order to extract the time participants needed to perform the tasks and participants' task success rates.

8.3. Results

This section presents an analysis of the data collected from blind users to investigate whether there was an improvement in users' experience through progressively improving the website design by removing the key type of problems they encounter. It begins with the presentation of participants performance (task success and task time) across the three website designs. Then it shows participants experience across the three website designs.

8.3.1. Participants' performance

The benefits of the participants' performance on each design solution were investigated. First, participants effectiveness (i.e. task success rates) was analysed. A total of 54 tasks were attempted in each website design. Table 34 shows participants' task success rate for the three website designs. Participants had very high success rates across all the different website designs, without any difference between the three website designs as assessed by Friedman test ($\chi^2(2) = 4.67$, $p = 0.097$).

Table 34. Participants' task success rates for each website design.

Website design	Task succeeded	Task failed
Structure only	96%	4%
Structure and feedback at the top of the page	94%	6%
Structure and in-place feedback	100%	0%

The second performance measure investigate was participants' efficiency (i.e. task time) using Friedman test on the participants' average task time. The analysis revealed that there was not any significant difference for participants' task time between the website designs ($\chi^2(2)$ = 3.13, p = 0.209). Table 35 shows participants' tasks mean time in seconds for each website design.

Table 35. Mean (SD) of participants' task time in seconds for each website design.

Website Design	**Mean Task Time in seconds (SD)**
Structure only	158.41 (73.69)
Structure and feedback at the top of the page	181.44 (99.36)
Structure and in-place feedback	125.87 (57.66)

8.3.2. Participants' experience

The benefits of each design solution on participants' experience were also investigated. An analysis of participants' answers to each question about their experience was conducted.

Perceived ease of use (Q1)

Participants rated how easy it was to use each website design. The rating scale was a 5-point Likert item, where 1 means "very easy" and 5 means "very difficult". Means and standard deviations for each website design are presented below. As it can be seen, participants found all website designs relatively easy to use.

Table 36. Mean (SD) of participants' perceived ease of use rating for each website design.

Website Design	**Mean Rating (SD)**
Structure only	2.28 (1.02)
Structure and feedback at the top of the page	2.00 (0.69)
Structure and in-place feedback	1.50 (0.51)

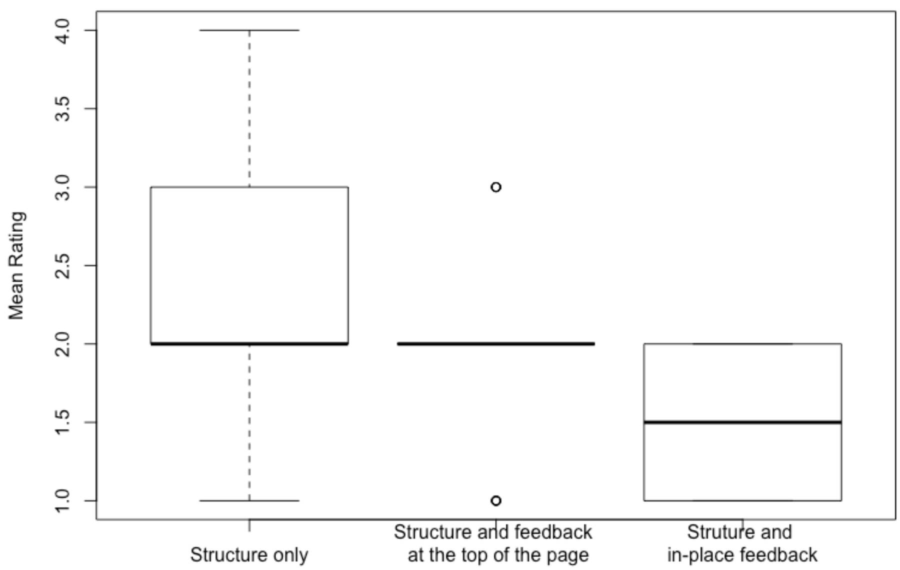

Figure 31. Boxplot showing the distribution of the perceived ease of use rating for each website design.

To investigate whether participants found easier to use one website design in comparison to the others, a Friedman test was conducted on participants' ease of use ratings. The analysis revealed a significant difference in the perceived difficulty between the website designs ($\chi^2(2)$ = 14, p = 0.001).

A post-hoc analysis with Wilcoxon signed-rank tests was conducted with Bonferroni correction, resulting in a significant level set at $p < 0.0167$. The analysis showed that participants found the *structure and in-place feedback* easier to use than *structure only* ($Z = -2.91$, $p = 0.004$, $r = -0.486$) and *structure and feedback at the top of the page* ($Z = -2.71$, $p = 0.007$, $r = -0.452$). The other comparison, *structure only* and *structure and feedback at the top of the page*, was not significantly different in terms of participants' ratings of ease of use ($Z = -1.41$, $p = 0.16$, $r = -0.234$).

Task completion confidence (Q2)

Participants felt very confident that they completed their tasks across all the three website designs, as can be seen in Table 37. To investigate whether participants felt more confident in one website design in comparison to the others, an analysis of their task completion confidence ratings was conducted.

Table 37. Mean (SD) of participants' task completion confidence rating for each website design.

Website design	**Mean task completion confidence (SD)**
Structure only	4.61 (0.61)
Structure and feedback at the top of the page	4.44 (0.62)
Structure and in-place feedback	4.67 (0.59)

A Friedman test revealed that there was not any significant difference between participants' confidence that they felt more confident completing the tasks more successfully in one website design in comparison to the others ($\chi^2(2) = 5.20$, $p = 0.074$).

How clear was what was happening in the page in response to users' actions (Q3)

An investigation of how clear was it for participants what was happening on the page in response to their actions on each design solution was performed. Means and standard deviations for each website design are presented below. The rating scale was a 5-point Likert item, where 1 means "not at all clear" and 5 means "very clear".

Table 38. Mean (SD) of participants' how clear was what was happening in the page in response to their actions rating for each website design.

Website Design	**Mean Rating (SD)**
Structure only	3.67 (1.09)
Structure and feedback at the top of the page	4.33 (0.59)
Structure and in-place feedback	4.72 (0.58)

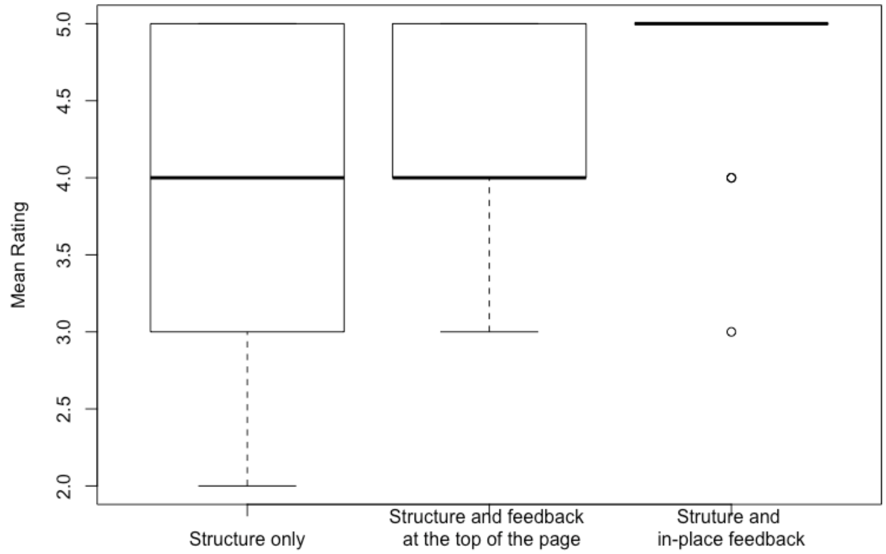

Figure 32. Boxplot showing the distribution for how clear was what was happening in the page in response to users' actions rating for each website design.

The analysis revealed that there was a significant difference between the three designs for how clear was it what was happening in the page in response to users' actions, as assessed via Friedman test, $\chi^2(2) = 20.15$, $p < 0.001$.

A post-hoc analysis with Wilcoxon signed-rank tests was conducted with Bonferroni correction, resulting in a significant level set at $p < 0.0167$. The analysis showed that participants felt that it was less clear happening in response to users' actions in *structure only* compared to *structure and feedback at the top of the page* ($Z = -2.81$, $p = 0.005$, $r = -0.468$) and to *structure and in-place feedback* ($Z = -3.13$, $p = 0.002$, $r = -0.522$). There was also a significant difference between the two-website design that provided feedback on users' actions, with participants founding clearer the website design with the feedback in-place (Z = -2.65, p = 0.008, r = -0.441).

Perceived workload (NASA TLX)

The perceived workload imposed by each design solution on participants was assessed. For workload, both the overall score of NASA-TLX and the scores of each of its subscale were calculated. Means and standard deviations for each website design are presented below.

Table 39. Mean (SD) of participants' NASA TLX overall workload score for each website design.

Website Design	Mean Rating (SD)
Structure only	7.20 (3.11)
Structure and feedback at the top of the page	6.79 (2.67)
Structure and in-place feedback	4.66 (2.17)

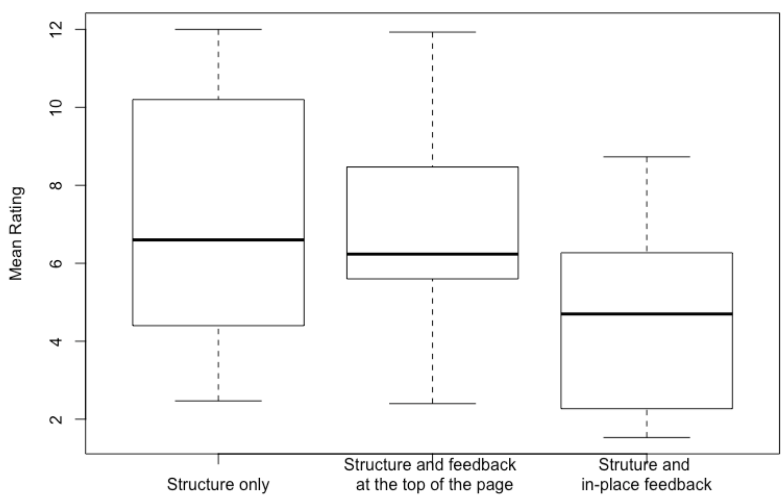

Figure 33. Boxplot showing the distribution of the overall workload rating for each website design.

A Friedman test on participants' overall NASA-TLX score revealed a significant difference between the websites designs, $\chi^2(2) = 14.11$, $p = 0.001$. A post-hoc analysis with Wilcoxon signed-rank tests was conducted with Bonferroni correction, resulting in a significant level set at $p < 0.0167$. The analysis showed that participants overall workload in *structure and in-place feedback* ($M = 4.66$, $SD = 2.17$) was lower than in *structure only* ($M = 7.20$, $SD = 3.11$), $Z = -3.20$, $p = 0.001$, $r = -0.534$, and *structure and feedback at the top of the page* ($M = 6.79$, $SD = 2.67$), $Z = -3.68$, $p < 0.001$, $r = -0.613$. The other comparison, *structure only* and *structure and feedback at the top of the page,* was not a significant difference in participants' mean overall NASA-TLX score, $Z = -0.74$, $p = 0.459$, $r = -0.124$.

The summary of the NASA-TLX subscale scores for the different website designs is shown in Table 40. In addition, an analysis of each NASA-TLX subscale score was performed (see Table 41).

Table 40. Mean (SD) of participants' NASA-TLX subscale scores for each website design.

	Structure only	**Structure and feedback at the top of the page**	**Structure and in-place feedback**
Mental Demand	23.89 (18.03)	25.72 (17.62)	21.61 (15.49)
Physical Demand	12.17 (15.11)	12.56 (16.95)	5.06 (6.04)
Temporal Demand	13.00 (13.34)	13.83 (13.30)	12.61 (9.70)
Performance	11.06 (9.71)	12.78 (9.27)	9.78 (7.70)
Effort	24.50 (20.63)	21.94 (16.99)	14.94 (12.88)
Frustration	23.33 (23.45)	15.00 (16.74)	5.83 (6.96)

Table 41. Analysis of the six NASA-TLX subscales (using Friedman tests) between the three website designs. *p < 0.05, **p < 0.01, ***p < 0.001.

NASA-TLX subscale	***$\chi^2(2)$***	**p**
Mental Demand	2.27	0.321
Physical Demand	3.46	0.178
Temporal Demand	1.13	0.569
Performance	1.10	0.576
Effort**	10.19	0.006**
Frustration**	12.25	0.002**

Post hoc comparison with Wilcoxon signed-rank tests was conducted with Bonferroni correction for effort and frustration subscales, resulting in a significant level set at $p < 0.0167$. The analysis showed that participants' effort and frustration scores were lower in *feedback in-place* compared to the other two designs. In more details, participants' effort score was lower in *structure and in-place feedback* in comparison to *structure only* ($Z = -3.05$, $p = 0.002$, $r = -0.509$) and *structure and feedback at the top of the page* ($Z = -3.01$, $p = 0.003$, $r = -0.501$). There was also a difference in participants' frustration score, with *structure and in-place feedback* having a lower score than *structure only* ($Z = -2.86$, $p = 0.004$, $r = -0.476$) and *structure and feedback at the top of the page* ($Z = -2.66$, $p = 0.008$, $r = -0.443$). There was no difference in the mean effort scores ($Z = -0.47$, $p = 0.637$, $r = -0.079$) and mean frustration scores ($Z = -2.14$, $p = 0.033$, $r = -0.356$) between *structure only* and *structure and feedback at the top of the page.*

Website design preference

To investigate whether participants preferred using one design over the others, a Friedman test was conducted on the participants designs preference rankings. The results of the analysis showed that there was a significant difference in participants' preferences between the designs, $\chi^2(2) = 28.78$, $p < 0.001$. Table 42 shows the frequency of ranks of the three website designs.

Table 42. Frequency of ranks for the three website designs.

Website design	**Rank 1st**	**Rank 2nd**	**Rank 3rd**
Structure only	-	3	15
Structure and feedback at the top of the page	-	15	3
Structure and in-place feedback	18	-	-

A post-hoc analysis with Wilcoxon signed-rank tests was conducted with Bonferroni correction, resulting in a significant level set at $p < 0.0167$. The analysis showed that *structure and in-place feedback* was more preferred than the *structure only* ($Z = -3.91$, $p < 0.001$, $r = -0.651$) and the *structure and feedback at the top of the page* ($Z = -3.91$, $p < 0.001$, $r = -0.651$). There was also a trend to a significant difference with *structure and feedback at the top of the page* ($Z = -1.89$, $p = 0.059$, $r = -0.314$) being more preferred than the *structure only*.

8.4. Discussion

The study presented in this chapter investigates the benefits of specific design solutions to prevalent problems on blind users' experience in a search and browse websites. Also, it checks the differences of providing simple design solutions to the problems with more complex design solutions that involve a combination of several design solutions. Table 43 below summarises the findings of the study.

Table 43. Summary of the results of the study.

Task Success	No difference between the three website designs.
Task Time	No difference between the three website designs.
Perceived Ease of use	Structure with feedback in-place perceived easier than just structure and structure and feedback at top.
Confidence	No difference between the different designs.
Clear	Structure with feedback in-place is clearer than the other two designs. Structure at top of the page is clearer than just structure.
Workload	Structure with feedback in-place is lower in workload than structure or feedback at top. Also, it required less effort and frustration than the other two designs.
Preference	Structure with feedback in-place was preferred more than the other two designs.

One of the most interesting findings was that participants were able to complete the tasks with very high success rates in all website designs. Even in the website design that no instant feedback was present, users succeeded over 90% of the time. While we cannot generalise that there is a difference between each of the website designs, the fact that each website has low rates of failures tends to indicate the importance of the structure in helping users to get an overview of the overall structure of the page content and navigating and finding particular information on the page. It provides further support that clear structure can improve participants performance, by not only making more efficient (T. Watanabe, 2009), but also more effective on the web. It also supports the recommendation of using headings to organise the page content by the accessibility guidelines (Caldwell et al., 2008). The benefits of the good structure are also supported by the comments made by participants during the sessions (see Table 44 below). In all website designs users praised the presence of structure:

Table 44. Participants' comments of how was it use each website design.

Comments	**Participant**
structure only	
"…nice structure, there were headings. It was easy to find the filters and the options were easy to select..."	P6
"...the filter titles were headings, so I could jump easy..."	P14
structure and feedback at the top of the page	
"…it is fundamentally easy. The headings are reasonable obvious. Everything works from top to bottom, it's quite simple…"	P8

"it was quite easy to follow when you understood the structure layout"	P4
structure and in-place feedback	
"Because the form fields were quite easy to navigate and the filters were headings. It gave you the information immediately when you select a filter which was useful. I found easy to navigate through the filters quickly."	P1
"...the headings were the key, well structure that's the important thing. The spoken feedback helps to reassure me that something happened, I got a sense that it worked..."	P6

Looking at the different user experience measures, there are some clear differences between the three website designs. Of particular interest, the structure with feedback in-place was perceived easier to use, made it clearer what was happening in the page in response to users' actions and imposed lower workload than the other two website designs. In contrast, there were very few differences between just adding feedback at the top of the page compared to just providing structure. Indeed, the feedback itself only has a detectable effect in terms of making it clearer for users of what is happening on the page, but it did not increase the perceived ease of use as one might expect. It is somewhat surprising that it did not have much improvement on users' experience, given that there were positive comments from participants about the feedback such as:

> *"when I click one of the checkboxes and it reload of the page and it included the message about the results and the filters" (P3)*

"..because when the screen refresh, it told you specifically what criteria have met. It told you what the results was, what the criteria were establish. You were aware what the results were, as soon as you actually refresh the screen" (P4)

"..when you jump at the top of the page you getting feedback straight away what you have already done…In your head, you immediate confirm that you click the correct ones." (P8)

This result tends to indicate that while the feedback on users' actions is important and provides the benefits we want of informing users what was the result of their actions, the extra effort users have to put in to traverse through the page again is overshadowing any benefits that might be present from the feedback. When looking at the specific workload components of NASA-TLX, the only detected difference between the designs, except of the expected effort, is that the structure with feedback in-place produced less frustration than the other two designs. The frustration users having to retrace their steps due to the refreshing of the page supports the idea that the benefits of feedback are reduced by the extra effort required. This idea is also supported by the comments relating to the structure page with feedback in-place. Many users talk about saving effort of returning at their lost position but also that they were able to proceed with their task without having to switch context:

"I like it because the feedback was concise, the feedback becomes separate from what you are actually doing, so it does not interfere with the process that you are doing to actually interact with the website.... What was great I am not rooting at the top of the page all the time, so I am getting this feedback so it is happening when I am

actually doing something. I am getting updating feedback and that made difference." (P8)

In this comment, the user is reflecting on the fact that they were able to continue their task and not be interfered with. They talk about how they did not have to root at the top of the page each time, a task itself as users need to retrace their steps back to where they were on the page, which means they have to temporarily postpone their existing task each time.

As one participant (P4) put it:

"Two reasons. One, it told you, it told you what the results of your selections was, but also it enables you to put all of your selections before, without having to refresh the screen. So, you did not need to go through the process again, you could go down the list and check them, but it told you as you went along what was happening so it was less intensive and onerous task."

The participant here talks about how they do not have to repeat the process, but also how they could "go down the list and check them", completing all the changes at once and monitoring the feedback.

The quantitative results, along with participants comments, tend to indicate that the extra effort that users are putting is not just because they have to traverse through the website. There is frustration that comes from repeatedly having to switch context. They have to temporarily postpone their primary task, engage in the traversal task to find where they were on the page and then re-engage with their primary task. This is a subtler understanding of the problems blind users have related to the extra effort. There is frustration related to having to navigate back to their lost position, but also frustration related to switching context repeatedly. Users need to remember not only their place on the page, but also where they were in their current task, and try to re-engage with the task in a meaningful way. However, even when there is good structure and understanding of what happened on the page, the frustration remains.

Of perhaps equal importance, users expressed a clear preference towards the structure with the in-place feedback. Given all the tangible benefits of users' experience and the strong preference by blind users for this website design, there is support for this being the recommended way to implement search and browsing websites.

The findings of the study showed that there is clear evidence of the immense importance of the page structure on users' experience, so much as task failure rates were almost disappeared. This result further supports the findings of T. Watanabe (2009) of the usefulness of headings on blind users' experience, in a study with much more blind participants, and the need to create websites with good structure. Therefore, web designers and developers can improve blind users' performance on the web by using headings to organise the page content. This result is quite important as it shows that structure is the dominant feature that should be addressed first on a website.

Even though no support was found that the common interactivity problems, extra effort due to the refreshing of the page and lack of feedback on users' actions, influence the effectiveness and efficiency of users on the web, it was found that it improves the overall user experience. However, when only the lack of feedback is addressed the benefits of users' experience are not much, except of making it clearer of what was happening on the page. Participants excessive effort seems to outdo all the benefits of the extra feedback. Blind users overall experience seems to be improved when a combination of key problems are addressed. When the extra effort due to the refreshing of the page accompanied with the lack of additional feedback on users' actions are addressed, it improves the overall users' experience. This result shows that simple design solutions (i.e. addressing the structure of the page only) help to improve specific usability measures. However, they are not enough to guarantee an improvement to the overall user experience. The cumulative effects of providing a combination of design solutions can provide a major improvement in the overall user experience. This also highlights that addressing these problems is quite complex because they influence each other.

Although the study provided us with a better understanding of how common interactivity problems influence the effectiveness, efficiency and the perceived usability of the website, as well as how specific design solutions can help to create a better user experience on the web, there are a few limitations that need to be considered. The tasks and websites used were not fully representing an exploratory search and browse, as users did not have to navigate between different pages, they did not have to interact with irrelevant content that may encounter in an exploratory search environment and the tasks were known-item searchers. This does not give a clear picture whether the findings are applicable to an exploratory search and threat the external validity of the results. In an exploratory search participants' cognitive load may be higher and as a result, the results of the task performance and users' experience could be different. Further research should be done to confirm whether findings of the benefits of users' experience through the different website designs apply to an exploratory search.

Another limitation of the study is that 5 out of the 18 participants took part in another study of this dissertation (study in Chapter 5), and one took part in the study in Chapter 4 as well. The participation in previous studies of this thesis can be a threat to the internal validity of the results. This limitation occurred, as it was difficult to recruit participants from such a small participant pool. It is not believed the recruitment of some participants that took part in previous studies had any impacts on the results. First, this study was conducted nine months after the study in Chapter 5. Also, the scope of this study was much different from the previous studies of this thesis. This study was focusing on users' experience in three different website designs, where participants did not perform a verbal protocol to elicit any problems they encountered on the website as in the previous studies of this thesis.

A limitation of the study that can lower the ecological validity is that participants did not use their own equipment. It was preferred not to ask participants to use their own equipment as this would require recruiting participants that have a laptop computer, which would have made the recruiting process even more challenging as not all blind users would have a laptop at home or at work. Participants were asked though to configure the equipment in their own preferences in order to match their usual setup. Also, participants first performed a task in practice website that data were not collected. Participants got familiar with the equipment they were using before performing the study tasks and any impacts of not using their own equipment were mitigated.

Another limitation of the study is that it collected data of multiple measures from the same participants. For example, participants' task success rate, task time, experience, workload and preference. Looking into differences between participants ratings on all these measures it may be considered as over-testing of the data. There is a possibility that there may have been a relationship between some of the measures and some tests were pushed into a significance. Ideally, to test whether there was a difference between the website designs on each measure, separate studies should have been conducted. However, due to pragmatic limitations of recruiting blind users, this was not preferred. Also, it is common in research with disabled users to collect data for more than one measure (Coyne & Nielsen, 2001; Disability Rights Commission, 2004).

8.5. Conclusions

This study set out to investigate how design solutions that progressively resolve the key problems blind users had in Chapter 5 influence users' experience on the web.

Based on the results, the most important design feature for improving blind users' effectiveness on the web is the page structure. If the page structure is properly implemented, it can be very beneficial for blind users as it greatly reduces the task failure on the web. Further, addressing the extra effort required due to the refreshing of the page can be of great help as users will not have to switch context, postpone their task to retrace their steps back to where they were on the page. The additional feedback on users' actions can provide more clarity about what is happening on the page but it would not be of much help if the page structure or the extra effort are not addressed first.

The study contributes to a better understanding of how specific design solutions can address interactivity problems to improve users experience on search and browse websites. Moreover, the study enhances our understanding that simple design solutions may address some of the problems blind users have but may not improve the overall experience. However, a combination of several design solutions can provide cumulative effects resulting in a major improvement in experience. This also shows that solving these problems is quite complex as the problems influence one another. Due to an important limitation of the tasks and website designs used in the study, a further research is required to confirm whether the findings are applicable to an exploratory search.

Chapter 9. Empirical confirmation study of the benefits of specific design solutions on blind users' experience in an exploratory search

9.1. Introduction

The previous study of this thesis investigated the benefits of specific design solutions to key problems on blind users' experience. The study provided a better understanding of how specific design solutions could address interactivity problems to improve users' experience. Also, it demonstrated that specific design solutions can address some of the problems, however, to improve the overall experience more complex design solutions required that use a combination of different design solutions.

There was a limitation in the previous study that makes it unclear whether the results can be applied in an exploratory search. During an exploratory search, participants will have to navigate between different pages but also encounter content that is not related to their task. The cognitive load of participants would be higher and that may cause the results found in the previous study not hold when users asked to do an exploratory search. To overcome this limitation, this confirmation study was set up that investigates whether the benefits found from the specific design solutions in the previous study of this thesis maintain when users are doing an exploratory search.

This study uses the same website designs as in the previous study, with the only difference of including two of the designs (*structure only* and *structure and in-place feedback)*. The *structure and feedback at the top of the page* design was not used in this study (see Chapter 7). The rationale of not using this website design is because it did not differ from the *structure only* in most the dependent variables compared, such as participants' task completion rate, task time, perceived ease of use, perceived workload and preference. Also, it was based on a non-standard approach that is not currently used by any website.

The unified definition of web accessibility proposed in Chapter 3 was used to devise the study. In relation to the key components of the unified definition, the study manipulates the website (design) only. The collected measures included qualities of usability, such as users' effectiveness, efficiency and satisfaction on each website design.

The study addresses the following research question:

- Do the benefits of specific design solutions on users' experience maintain in an exploratory search?

9.2. Method

9.2.1. Study Design

A repeated measures within-participants design was used, with the design as the independent variable with two levels (*structure only* and *structure and in-place feedback*). The dependent variables were qualities of usability, such as participants' task success rate, task time, ease of use and perceived workload.

9.2.2. Participants

Twenty blind screen reader participants took part in the study. Sixteen of the participants were men and four were women. Ages ranged from 25 to 72 (M = 51.1, SD = 16.1). Ten of the participants were congenitally blind while the remaining ten lost their sight between the ages of 3 and 42. Due to the limited participants pool, 12 out of 20 participants were participants that took part in the previous study of this thesis (see Chapter 8). Of the 12 participants, two of them took part in the study in Chapter 5 as well. From the remaining eight participants, three of them took part in the study in Chapter 5, of which two of them also took part in the study in Chapter 4.

Participants rated their experience and expertise on the web using a five-point Likert item, where 1 means “very low” and 5 means “very good”. The average rating of web experience was 4.3 (*SD* = 0.7), whereas the average rating of expertise was 4.0 (*SD* = 0.7).

All participants used screen readers to access computers and the web from home and work. Fifteen participants used JAWS (running on Windows OS), three used NVDA (running on Windows OS) and two used VoiceOver (running on Mac OS). The JAWS version used varied from JAWS 15.0 to JAWS 17.0 (the latter being the latest version of JAWS when the study was conducted). Participants who used NVDA used the latest version 2016.2.1. Participants who used VoiceOver used the latest version that comes with Mac OS El Capitan (the latest version of Mac OS when the study was conducted). Participants were asked to rate their experience and expertise using screen readers on a five-point Likert item, where 1 means “very low” and 5 means “very good”. The average rating for experience and expertise using screen readers was 4.5 (*SD* = 0.7) and 3.9 (*SD* = 0.6), respectively.

Eighteen participants used Windows and two participants used Mac OS. The majority of the participants who used Windows reported Internet Explorer as their primary browser and all the participants who used Mac OS reported Safari as their primary browser.

9.2.3. Equipment and Material

For participants who used the Windows OS, the study was conducted using a desktop computer running Windows 10 with speakers and keyboard. For participants who use the Mac OS, the study was conducted using a MacBook Pro running the El Capitan OS with speakers. In addition, participants were able to choose the screen reader software they were most familiar with, for example, JAWS, NVDA or used the VoiceOver version that comes with El Capitan OS on Mac. The screen reader software that participants used was already declared during the recruitment process and all installation of the software was already arranged properly before the arrival of the participants to match their home or work environment.

Participants did not use their own equipment as I wanted to ensure that the equipment was in running order before the arrival of the participant. Also, the sessions were recorded using Morae 3.1 on Windows and ScreenFlow 4.0.3 on Mac OS, that was preinstalled on the computers used in the study. These recordings included audio and screen activity.

When participants completed all tasks on each website they were asked to complete a questionnaire about the website using a 5-point Likert items:

- Q1: How easy or difficult did you find the website to use?
- Q2: How confident or not confident are you that you completed the tasks successfully?
- Q3: How clear or not clear was it to you, what was happening in the page in response to your actions?

Then, participants were asked about how was it to use each website design. Participants were also asked to complete the NASA TLX, a subjective workload assessment questionnaire (Hart & Staveland, 1988).

9.2.4. Websites and Tasks

The two website designs used in this study were the *structure only* and *structure and in-place feedback* from the study in Chapter 8. To achieve high ecological validity a setting that matches as close as possible a real-life situation was implemented. The websites created had content from two commercial shopping websites. This included the products of five different categories from a furniture website (Habitat) and a technology website (Currys). For the furniture content website, data for the beds, dining tables, footstools, sofas and wardrobes were extracted. For the technology content websites, data for laptops, tablets, phones, televisions and headphones were extracted. The import.io tool was used to support the extraction of data from the websites. All references to the initial websites were removed from the data collected.

To make the websites as close as possible to a real-life website, each website had a homepage containing navigation options as well as featuring the top picks from each product category. At the header and footer of the page, information about the website was provided, such as store locator, contact us, links to social media. Figure 34 and Figure 35 show an example of one of the website designs.

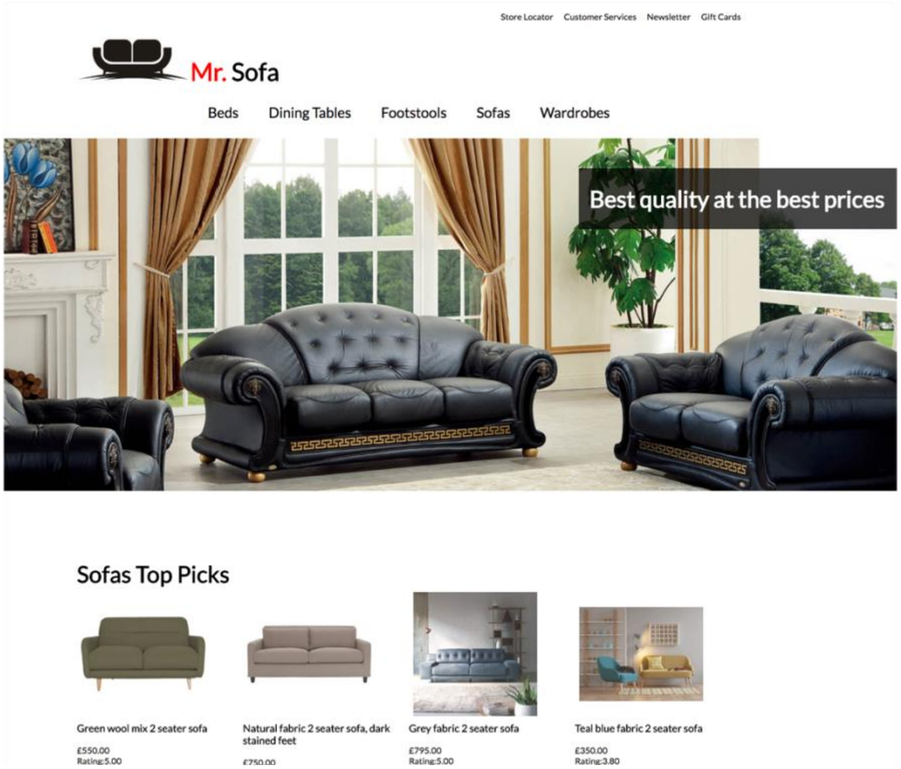

Figure 34. Example of the homepage of a website design.

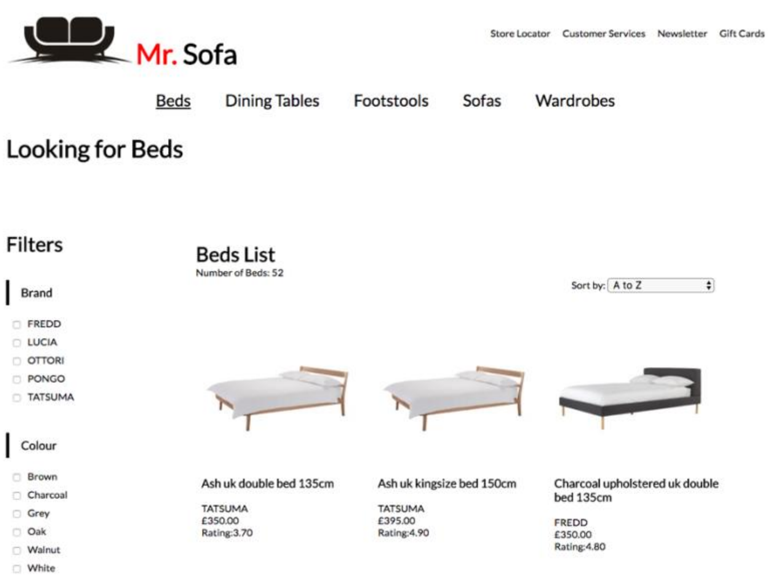

Figure 35. Example of the filters and browsing page content of a website design.

On each website, participants were given an introductory scenario and then they were asked to perform four exploratory search tasks. The scenarios and tasks for each website were:

Mr. Sofa

Scenario: You have just moved into a new house and you want to buy new furniture. A friend suggested using Mr. Sofa, a company that offers good quality for reasonable prices. You decided to visit their website and check online their furniture.

Tasks:

- Task 1: Find the price of the cheapest King Size Bed.
- Task 2: Find the highest rated glass dining table, that cost less than £300 and fits either 4 or 6 people.
- Task 3: Find the dimensions of the cheapest 3-seat sofa, that its material is fabric, its colour is grey and it’s rated with more than 3 stars.
- Task 4: Find the cheapest of the highest rated wardrobes, that its colour is White or Mirrored, its material is Lacquered, and has either 2 or 3 doors.

The Gadget Shop

Scenario: You want to buy new techs. A friend suggested that you could find what you are looking for at The Gadget Shop, an online shop for technology items. You decided to visit The Gadget Shop website first before you go to the shop.

Tasks:

- Task 1: Find the price of the cheapest TV that its screen size is between 30 to 40 inches.

- Task 2: Find the highest rated in-ear headphone, that costs less than £100 and its brand is either Beats or Bose.
- Task 3: Find the screen size of the cheapest phone that its brand is Apple, its internal memory is 64 GB, it's rated with more than 3 stars and runs iOS operating System.
- Task 4: Find the cheapest of the highest rated laptops, that its operating system is Windows, has 512GB or 1TB of storage and its memory is either 4GB or 8GB.

In preparation of the evaluation, the tasks were first undertaken using JAWS and NVDA on Windows and VoiceOver on Mac OS, to check that it is possible for screen reader users to be able to complete the tasks.

9.2.5. Procedure

The study took place in the Interaction Laboratory at the Department of Computer Science of the University of York. Participants were briefed about the study and were asked to sign an informed consent form. To avoid any conflicts between the technology and participants' preferences, participants were asked which screen reader and browser would like to use. Then, they were given the option to adjust the sound and related software to their preference in order to match their usual setup.

Participants first performed three tasks in a training website, which was not used in the analysis of the data. The present study was piloted with one blind participant, whose data were not used in the analysis of the study.

In total, four websites were created to cover each combination of design and content. The order that participants evaluated each website was counterbalanced. The tasks order on each website was not counterbalanced as the order of the tasks was of increasing difficulty, based on the steps required to perform the tasks.

After participants had completed all the tasks in all websites, they were asked to complete a demographic questionnaire. Then participants were debriefed about the study and I answered their questions. Finally, any information that was necessary for the compensation of participants' time was collected.

9.2.6. Data Analysis

The video recordings of each participant were reviewed, to extract the time participants needed to perform the tasks and participants' task success rates.

9.3. Results

This section presents the analysis of the data collected from blind users regarding the benefits of each design solution to improve the effectiveness and efficiency of users and the perceived usability of a website. This will help us to assess whether the benefits found in the previous study maintain in an exploratory search. First, it presents the benefits of participants performance (effectiveness and efficiency) for the two design solutions. Then it focuses on the benefits of participants' experience across the designs.

9.3.1. Participants' performance

The first performance characteristic investigated was participants effectiveness (task success rates). A total of 80 tasks were attempted on each website design. Table 45 shows participants' task success rate for the two designs, with the percentage of the tasks succeeded and failed for each design.

Table 45. Participants' task success rates for each website design.

Website design	**Task succeeded**	**Task failed**
Structure only	95.0%	5.0%
Structure and in-place feedback	97.5%	2.5%

Participants had very high success rates on both websites, without any significant difference in participants' success rates between the two designs, as assessed by Wilcoxon signed-rank test (Z = -1.00, p = 0.317, r = -0.158).

When considering participants' effectiveness, an analysis of participants average task time on each website design was performed. Table 46 shows participants' tasks mean time in seconds for each website design. The analysis revealed that there was significant difference for participants' task time required between the two website designs, with participants requiring more time in *structure only* (M = 328.85, SD = 153.13) than in *structure and in-place feedback* (M = 253.95, SD = 100.09), as assessed by Wilcoxon signed-rank test (Z = -3.02, p = 0.002, r = -0.478).

Table 46. Mean (SD) of participants' task time in seconds for each website design.

Website Design	**Mean Task Time in seconds (SD)**
Structure only	328.85 (153.13)
Structure and in-place feedback	253.95 (100.09)

9.3.2. Participants' experience

The benefits of participants' experience were investigated. This included an analysis of participants answers on each question about their experience with the website design as well as their workload.

Perceived ease of use (Q1)

Participants rated how easy was it use each website design using a 5-point Likert item scale, where 1 means "very easy" and 5 means "very difficult". Means and standard deviations for each website design are presented below.

Table 47. Mean (SD) of participants' perceived ease of use rating for each website design.

Website Design	**Mean (SD)**
Structure only	2.60 (1.00)
Structure and in-place feedback	1.55 (0.51)

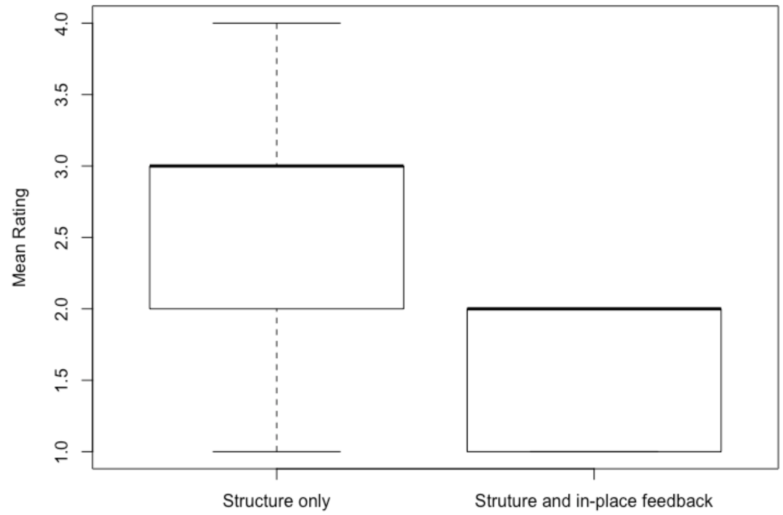

Figure 36. Boxplot showing the distribution of participants' perceived ease of use rating for each website design.

As can be seen from the figure above, participants' average ease of use ratings were relatively low in both designs, meaning they found the websites easy to use. Further analysis was conducted to check whether there was a significant difference in participants ease of use ratings between the two designs. The analysis revealed a significant difference in the perceived ease of use ratings between the two website designs, with participants founding the *structure and in-place feedback* (*M = 2.60, SD = 1.00*) easier to use compared to the *structure only* (*M = 1.55, SD = 0.51*), as assessed by Wilcoxon signed-rank test ($Z = -3.54$, $p < 0.001$, $r = -0.559$).

Task completion confidence (Q2)

The mean ratings of participants' task completion confidence were very high in both designs (see Table 48). Moreover, an analysis of participants task completion confidence ratings between the two designs was performed to investigate whether participants felt more confident completing their tasks in one website design in comparison to the other. The analysis showed that participants did not feel more confident about their task completion in one website design in comparison to the other, as assessed by Wilcoxon signed-rank (Z = -1.13, p = 0.257, r = -0.179).

Table 48. Mean (SD) of participants' task completion confidence rating for each website design.

Website design	**Mean task completion confidence (SD)**
Structure only	4.40 (0.75)
Structure and in-place feedback	4.55 (0.51)

How clear was what was happening in the page in response to users' actions (Q3)

An investigation of how clear was it for participants what was happening on the page in response to their actions on each design solution was performed. Means and standard deviations for each website design are presented below for participants ratings on how clear was it what was happening on the page in response to their actions was much higher in the design that included the instant feedback.

Table 49. Mean (SD) of participants' how clear was what was happening in the page in response to their actions rating for each website design.

Website Design	**Mean (SD)**
Structure only	3.10 (1.12)
Structure and in-place feedback	4.70 (0.47)

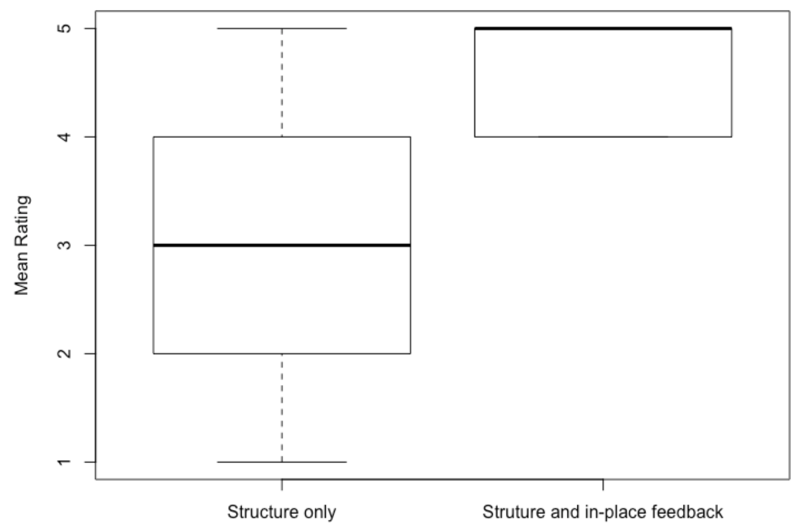

Figure 37. Boxplot showing the distribution of participants rating for how clear what was happening in the page was in response to their actions for each website design.

To check whether the difference was significant or not, a Wilcoxon signed-rank test was performed. The analysis revealed a significant difference between the website designs for how clear what was happening in the page was in response to users' actions, with *structure only* (*M = 3.10, SD = 1.12*) being less clear than *structure and in-place feedback* (*M = 4.70, SD = 0.47*), $Z = -3.58$, $p < 0.001$, $r = -0.566$.

Perceived workload (NASA TLX)

To investigate whether one design imposes extra workload on participants, both the overall score of NASA-TLX and the sub-scale scores were calculated. Means and standard deviations for each website design are presented below for the overall NASA-TLX score.

Table 50. Mean (SD) of participants' NASA TLX overall workload score for each website design.

Website Design	**Mean (SD)**
Structure only	9.95 (3.27)
Structure and in-place feedback	5.46 (2.00)

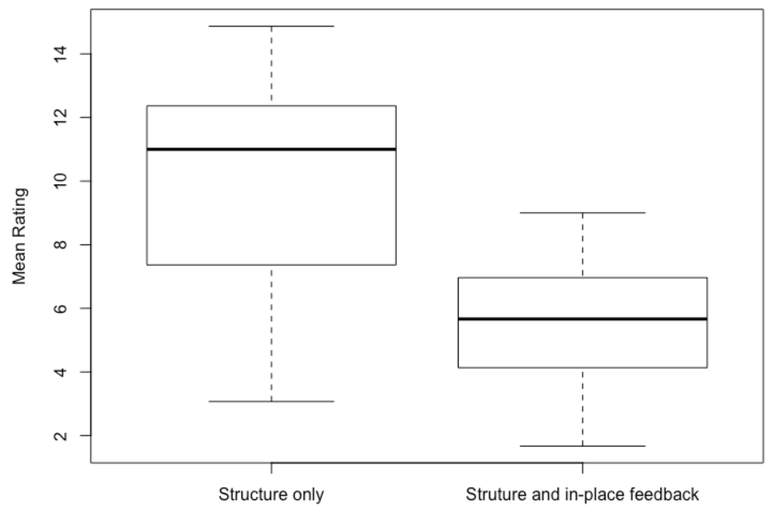

Figure 38. Boxplot showing the distribution of the overall workload rating for each website design.

A Wilcoxon signed-rank test revealed a significant difference in the overall workload score between the two designs, Z = -3.88, p < 0.001, r = -0.614. The overall workload score in *structure page* (M = 9.95, SD = 3.27) was higher than in *structure and in-place feedback* (M = 5.46, SD = 2.00).

Also, an analysis of the six NASA-TLX subscale scores was conducted. The summary and analysis of the NASA-TLX subscale scores for the two website designs are shown in Table 51.

Table 51. Mean (SD) and analysis (using Wilcoxon signed-rank tests) of NASA-TLX subscale scores for the two website designs. *p < 0.05, **p < 0.01, ***p < 0.001.

NASA-TLX subscale	Structure only	Structure and in-place feedback	Z	p	r
Mental Demand*	37.00 (21.36)	25.25 (14.83)	-2.56	0.010	-0.405
Physical Demand**	17.90 (21.26)	8.05 (10.08)	-3.11	0.002	-0.492
Temporal Demand*	15.45 (14.17)	11.40 (9.65)	-2.11	0.035	-0.334
Performance	8.75 (8.20)	9.85 (8.07)	-1.31	0.192	-0.206
Effort**	34.50 (19.92)	20.30 (14.83)	-3.11	0.002	-0.491
Frustration***	35.65 (27.33)	7.05 (7.47)	-3.51	< 0.000	-0.554

Wilcoxon signed-rank tests showed that there was a significant difference on five workload sub-scales (see Table 51). Participants' mental demand, physical demand, temporal demand, effort and frustration scores were lower in *structure and in-place feedback* compared to the *structure only.*

9.3.3. Investigation of whether the participation in the previous study had any impact on the results

Due to the limited participant pool, some of the participants took part in multiple studies. This is a study limitation, as it may have had an impact on the results and requires examination.

To investigate whether the participation in the previous study has had any effect on participants' performance and overall workload, an analysis between the two user groups was conducted. The user group of participants that took part in the previous study is referred as "repeated participation", whereas the user group of participants that did not take part in the previous study is referred as "non-repeated participation". Means and standard deviations for participants' performance and overall workload per user group are presented below.

Table 52. Mean (SD) of participants' performance and overall workload for the two website designs by user group.

	structure only		**structure and in-place feedback**	
	repeated participation	non-repeated participation	repeated participation	non-repeated participation
Task success rate	95.8%	93.8%	97.9%	96.9%
Task Time in seconds	332.23 (156.06)	323.78 (159.13)	252.19 (95.32)	256.59 (113.59)
Overall NASA-TLX score	10.32 (3.13)	9.40 (3.61)	5.62 (1.82)	5.23 (2.36)

As can be seen from the table above, the task success rates of the two user groups were very similar. Thus, only an analysis of participants' efficiency (task time in seconds) was performed in regard to participants' performance measures. The most appropriate test to compare the mean differences between groups that have been split into two factors, where one factor is within-participants and the other factor is between-participants is the two-way mixed ANOVA. The within-participants' factor is the website design that has two levels (structure only and structure and in-place feedback) and the between-participants' factor is the two user groups (repeated participation group and non-repeated participation). However, the data of effectiveness violate one of the assumptions of the test, as they are not normally distributed. As there is not any non-parametric test for a two-way ANOVA, the analysis was carried out as planned. The figures below show the distribution of the data for participants' task time and overall workload.

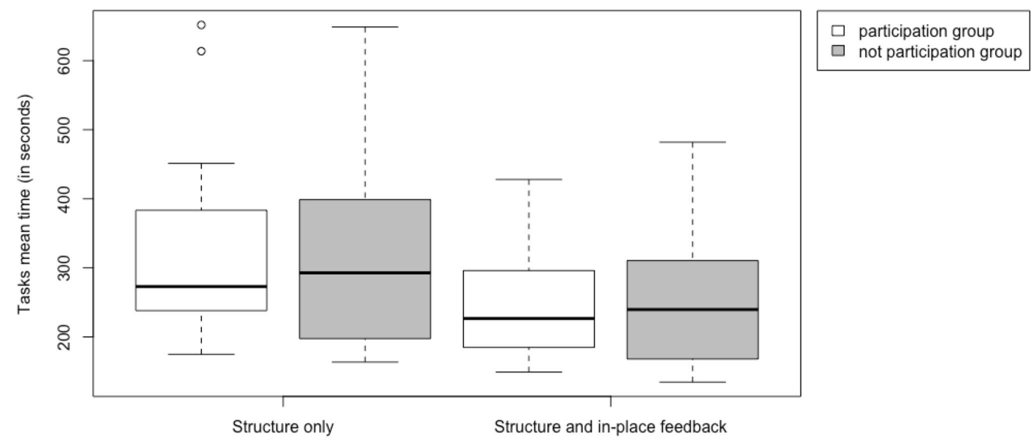

Figure 39. Boxplot showing the distribution of the tasks' mean time (in seconds) by user group.

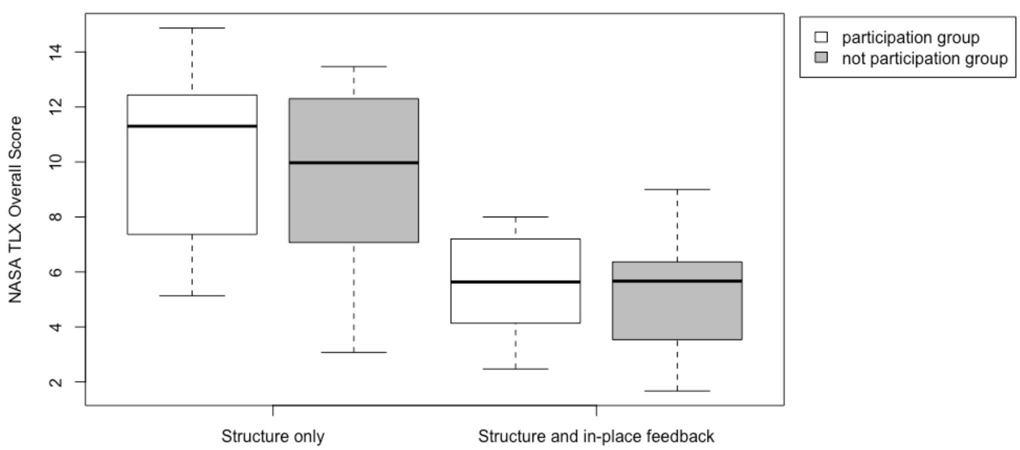

Figure 40. Boxplot showing the distribution of NASA TLX overall score by website design and user group.

Table 53 shows the results of the two-way mixed ANOVA on participants' task time and overall NASA TLX workload.

Table 53. Main effects and interaction effect of the two-way mixed ANOVAs.

	Main effect website design			**Main effect user group**			**Interaction effect**		
	F (1,18)	**p**	$\eta^2_{partial}$	**F (1,18)**	**p**	$\eta^2_{partial}$	**F (1,18)**	**p**	$\eta^2_{partial}$
Task Time	7.31	0.015	0.289	0.00	0.971	0.000	0.06	0.816	0.003
Overall	56.24	0.000	0.758	0.343	0.565	0.019	0.19	0.665	0.011

As can be seen from the table above, the participation to the previous study of the thesis did not have any impact on participants' task time or their overall workload. The main effects of the website design confirm the findings between the two website designs in subsections 9.3.1 and 0.

9.4. Discussion

The study presented in this chapter investigates the benefits of users' experience in an exploratory search by addressing key problems through specific design solutions. This is a confirmative experimental study with more ecological tasks and setting to assure the benefits found on users' experience in the previous study (Chapter 8) maintain. Table 54 below summarises the findings of the study.

Table 54. Summary of the results of the confirmation study.

Task Success	No difference between the website designs.
Task Time	Participants were more efficient in the structure with in-place feedback than in the structure only.
Perceive Ease of use	Structure and in-place feedback perceived easier to use than structure only.
Confidence	No difference between the website designs.
Clear	Structure and in-place feedback is clearer than structure only.
Workload	Structure and in-place feedback is lower in workload than structure only. Also, it was less mentally, physically and temporally demanding, less frustrating and required less effort.

The findings of this study in regard to participants effectiveness confirm the results of the previous study. Participants were able to achieve very high success rates on both website designs. The success rates found in this study are much higher compared to the ones found in studies in the literature (Disability Rights Commission, 2004; André Pimenta Freire, 2012; Petrie & Kheir, 2007) and in the study in Chapter 5. The fact that each website design had very low rates of failures indicate the importance of the structure in helping users getting an overview of the overall structure of the page content, navigating and finding particular information on the page.

Previous research (T. Watanabe, 2009), although with a very small number of blind users, showed that organising the page content with headings can benefit blind users' efficiency on the web. The results of the present study further expand the benefits of good structure on blind users' experience, as it also benefits their effectiveness. This result can also be used as strong empirical evidence to support the recommendations of the accessibility guidelines of using heading to organize the page content (Caldwell et al., 2008).

The benefits of the good structure are also supported by the comments participants made during the sessions (see Table 55):

Table 55. Participants' comments of how was it use each website design.

Comment	**Participant**
structure page	
"I liked the way it was structured, it had the navigation to navigate to the different sections, like laptops, phones, earphones, that made it very easy to go to the different sections. It was very easy laid out. It did not have too many graphics and It was very fast."	P1
"I liked the layout, it was easy enough to know where you were."	P5
"Structured, consistent, logical, it was easy to get straight to where I wanted."	P8
structure and in-place feedback	

"I like the structure, it had some continuity in flow to it. The feedback it told you where you were, it gave you a view that actually made you feel reasonably confident that you selected the right things."	P7
"I like the clarity of it, the way it was structured, it was quite easy… it had a nice hierarchy it had all the headings first... There was a nice structure there."	P15
"… headings, specifically I liked the fact that you left in the same place when it refreshed, that's such a big difference. And it was the big thing that I liked about it was the narrative of the what you are getting back, you get the story of what filters you put on. The feedback so you do not have to go hunt for it, so you are aware that you got the right feedback"	P16
"I liked how it was nice laid out, I like the how the filters worked, I found it quite easy to use"	P20

A usability measure that was different between the two website designs was participants' efficiency. Participants were more efficient performing their tasks in the website design with feedback in-place compared to the structure only. This result differs from the findings of the previous study. As users do not have to switch context each time they select a filtering option, which required them to retrace their steps to where they were on the page, it likely means that cumulatively there is time-saving as users browse in an exploratory search.

Looking into other user experiences measures, the findings of this study confirm the existing findings of the previous study in regard to the structure with feedback in-place design being easier to use and clearer about what is happening in the page in response to users' actions compared to the structure only.

What is of particular interest is the result of the workload. This study confirms the existing findings that the structure with feedback in-place requires less effort and less frustration than the structure only, but it also strengths the results by providing new knowledge. It was also found that structure with feedback in-place was perceived less mental demanding, less physically demanding and less temporal demanding than structure only. Participants cognitive effort was much higher as they had to switch context each time a filtering option was selected. Participants had to remember where they were on the page and navigate there, which seems to add on their already task cognitive load. In contrast, on the website design with in-place feedback participants were focus only on their task. The tasks being more mentally demanding are also supported by some of the participants' comments relating to the structure only design:

"For the most parts, it was easy to navigate, and it was well laid out. It was slightly disorienting you when you press spacebar on any of the filters it took you at the top of the page, that was a bit disorientating..." (P1)

"When you check a checkbox, it did not remain where it was…you are not sure whether it accepted it or which ones you have to check, you had to remember what you check and go find them again in the whole list again, you have to find the list every time." (P10)

In these two comments, P1 talks about the getting slightly disorienting due to the refreshing of the page, whereas P10 talks that they had to remember what they check to go and find it again in the list. As users need to remember where they were on the page to traverse back to that position it can impose higher cognitive workload on participants, as they do not only have to remember what they have to do on their task but they also have to remember where they were on the page before the page refreshed.

The difference in participants' physical effort has to do with the more actions that participants needed to make to perform the tasks. As one participant put it:

> *"Although it was consistent, it was usable, but because of the fact that when you check the checkbox it did not stay, you have to use the headings all the time to get back to where you needed to be so that took a lot more work that it should have done. It was difficult, that's what made it difficult." (P11)*

The participant here mentions that it "took a lot of more work that it should have". Due to the refreshing of the page participants had to do more steps that normally would need.

The extra workload of participants on this website design, as in the previous study, is not just because they have to traverse through the website. Participants had to postpone their primary task, remember where they were on the page and navigate to that section. This seems to be made participants feel more time pressure to complete the task as they had to quickly find where they were on the page and then continued their primary task. This result confirms the workload findings of the previous study and further strengthens them with new insights. It provides a clearer understanding of the benefits of specific design solutions on users' workload in an exploratory task.

Interestingly, the difference in participants' performance (i.e. efficiency) was not consistent with the performance measure of the NASA TLX questionnaire. Participants were more efficient in completing their tasks in website design with feedback in-place compared to the structure only. However, no support was found on the performance measure from NASA TLX questionnaire that the two website designs differ. A possible explanation for this might be the way the performance measure is phrased in the NASA TLX questionnaire: "How successful do you think you were in accomplishing the goals of the task set by the experimenter (or yourself)? How satisfied were you with your performance in accomplishing these goals?" (Hart & Staveland, 1988). The way the performance question is phrased is asking users to assess two different measures under one question (how successful and how satisfied). By asking participants how successful they think they were in accomplishing their goals set by the evaluator, participants may think about their effectiveness of completing the tasks, in which no difference was found between the two website designs. Also, the way the question is phrased is similar to the question about participants' confidence in completing their tasks successfully, in which again no difference was found between the two website designs.

The results of the present study confirm most of the findings of the study in Chapter 8 in an exploratory search. It confirms that the findings of effectiveness, perceived ease of use, how clear what was happening in the page in response to users' actions was and participants' overall workload are maintained in an exploratory search. Furthermore, it provides new knowledge in regard to participants efficiency, mental demand and physical demand workload measures. In this study, it was found that there was also an improvement in participants efficiency in the structure with in-place feedback website design. Participants required more time to perform the tasks when the page refreshes each time users selected a filtering option as they have to switch context. This also seems to influence participants’ mental load and physical effort when users are doing an exploratory search. Participants context switching imposed a greater cognitive load on participants but also required them to do more steps to progress with their task.

Taken together, this study confirms the results of the previous study of this thesis (Chapter 8), with the page structure being the dominant design feature of immense importance on blind users’ experience on the web. If the page is designed with appropriate headings to provide content structure, then blind users’ effectiveness will be enhanced. However, what was interesting is that the page structure needs to be accompanied with several other design solutions in order to provide a major improvement in the overall user experience. This result shows that solving the problems blind users have is quite a complex task as the problems relate to each other.

This study shows that there is a need for more empirical evidence-based investigations of design solutions on the problems blind users have on the web. Thus, design solutions that are accompanied by empirical evidence how they benefit users’ experience can be proposed.

There are a few limitations in the study that need to be considered. The study did not include the question about participants’ preference. As the results of the previous study (Chapter 8) showed that all participants ranked the *structure and in-place feedback* as their most preferred one, it was chosen not to include this question. However, reflecting back to the study design, it would have been more beneficial, particularly for comparison reasons, to had included this question in this study as well.

Another limitation of this study is that it primarily focused on measuring users’ experience and did not consider the collection of measures like participants’ keystrokes. It would have been interesting to check whether participants navigation strategies changed when they were navigating through the different website designs, and keystrokes would have been provided valuable information to this.

The most important limitation of this study is that half of the participants also took part in the previous study of this thesis (see Chapter 8), which can threaten the internal validity of the results. Although this was a limitation in other studies of this thesis, in this study, it should be given more importance as this was a confirmative study of the results of the previous study of this thesis. In this particular case, even though there was a 5-month gap between the two studies, it could have still impacted the results. Looking through into the analysis between the two user groups, there was not any difference between the two user groups on their task time or perceived workload. This limitation, however, has an important learning point in the field of human-computer interaction, when conducting studies with small participant pools. More consideration should be given when designing studies with participants that took part in other studies. For example, researchers should have a considerable time gap between studies that participants can take place. Recruit the same participants in studies of different nature. It is also very important for the researchers to check whether there were any differences between the participants that took part in the previous studies and the participants that did not.

The current study was conducted with participants that did not use their own computers, which lower the ecological validity. Although it would have been better if participants used their own equipment, it was preferred not to, as that would have made the recruiting process even more difficult. I would have to limit the participant pool to ones that had a laptop computer at home or at work. To alleviate any impacts, participants were given the option to configure the equipment to their own preferences to match their usual setup and first performed a practice task to get familiar with the equipment.

The study collected data of multiple measures from the same participants and looked at all these things in one study. There is a possibility of over-testing, as it might have been a relationship between some of the measures that push a test into a significance. Although it would have been better if there was only one primary dependent variable, due to pragmatic limitations of recruiting blind users, this was not preferred.

9.5. Conclusions

The study presented in this chapter is a confirmative investigation of the benefits of different design solutions on blind users' experience on the web by doing an exploratory search. The websites and tasks used in the present study were of higher ecological validity, as they represent real-life situations. Users had to navigate between different pages, encountered information that may not be relevant to their task and performed an exploratory search task.

The results indicate that the benefits of addressing the key problems through specific design solutions are also maintained in an exploratory search, but they also strengthen the results with new insights. Starting from the maintained benefits, the study confirms participants high success rates, which shows the immense importance of the page structure for improving blind users' effectiveness on the web. If the problems in relation to the structure of the page are properly addressed, then users' task failure rates will greatly reduce. Moreover, this shows the importance of the structure of the page not being neglected during the design of websites and it should be given higher priority when designers and developers conduct accessibility fixes on their websites. Looking into other user experience measures the study corroborates the benefits of the structure with in-place feedback on blind users' experience. Participants found it easier to use and it was clearer what was happening on the page in response to their actions.

Looking into new the new insights provided by the study, it shows that participants were more efficient in the website design with feedback in-place. Moreover, their overall workload was decreased, confirming the results of the previous study, but also their mental demand and physical demand. This result provides new knowledge of the benefits of these specific website design solutions on blind users' experience. Users cognitive load would be higher in an exploratory search as they will have to go and navigate through different pages, they will try to understand how the page content is structured and go through unnecessary information. It is important to design solutions that mitigate users' workload and create a better user experience.

This study confirms that the benefits of specific design solutions are maintained in an exploratory search. In addition, it shows that to improve the overall blind users' experience on the web is quite a complex task, as it requires a combination of different design solutions that can work together to provide an improvement in their experience. The cumulative effects of combining different design solutions can provide a major improvement in the overall user experience.

There is a need for further studies like this, in order to suggest design solutions to the problems blind users have on the web that are accompanied by empirical evidence of their benefits on users' experience. Researchers and designers should look further into areas that impose high workload on users, as simple design solutions such as the good page structure may not be able to address them. An area that worth further investigation to test is the benefits of specific design solutions to the information overload and the issues with the order users perceive the content on the page. It is unclear whether users will still have these problems when the pages are well structured, as they influence each other.

Chapter 10. Overall Discussion and Conclusions

10.1. Overall Discussion

This thesis work contributes to the field of accessibility by providing further insight into the problems blind users have on the web through a comparison and contrast of blind and sighted users' problems. In addition, it tested specific design solutions for some of the most prevalent problems blind users have and it provides a deeper understanding of how specific design solutions can benefit blind users' experience on the web.

The work demonstrated that the problems blind and sighted users have on the web largely differ. Without question, problems related to the physical presentation of the page content are distinct to sighted users. What is interesting is that only blind users had problems with the technology they were using. This shows that there are still mismatches between the assistive technologies blind users use, the browsers and the websites. Even that the two user groups encountered many similar types of content, information architecture and interactivity problems, the characteristics of the problems were very different. Also, there were problem types distinct to blind users, such as the problems with the page structure and the lack of feedback on users' actions. Many of the issues blind users had were due to poor technical implementation. For example, interactive elements without associated labels, links without accessible descriptions, images without alternative content, functionality not working. For many of these problems, the accessibility guidelines provide techniques that can help designers and developers to address them.

However, many of the key problems blind users encountered were due to the high cognitive workload and mostly due to the sequential processing they have to do. In contrast to sighted users, blind users do not have a full overview of the page structure instantly. In addition to that, if the page is not well-structured users may have difficulties to understand the structure of the page content as well as navigate and find particular information on the page. Even when they can navigate around the page using the page headings they still have to sequentially process the page content and thus does not disregard content like advertisements or other irrelevant material.

Indeed, the page structure is of immense importance for improving blind users experience, as it can help them understand how the page content is structure and navigate easier through the page content. However, simple design solutions like this will not be enough to guarantee a major improvement in the overall experience of users. Websites with good structure can still impose high workload on users if other problems are not addressed. One of the main reasons for the high workload was the context switch users had to do each time the page was refreshing. Blind users had to postpone their primary task, remember the position they were on the page, navigate back to that position, and then they re-engage with their primary task. This issue is quite important as it can overshadow all the benefits users have by addressing other problems on the page (i.e. lack of feedback on users' actions). Although users found the pages with the additional feedback clearer of what was happening on the page, all the benefits of feedback were mitigated due to the context switch users had to do. This provides a deeper understanding of the benefits of specific design solutions to the problems blind users have. Simple design solutions to the problems may improve specific user experience measures, however, to provide a major improvement in users experience a combination of several design solutions is required. This result also shows that addressing these problems is a complicated task, as it is not only about providing good structure, such as using headings to structure and organise the page content but also looking for areas that impose high workload on users and test if and how design solutions can reduce it.

To create a better accessible user experience on the web for blind uses there is a need to test more design solutions that will be empirically supported for future implementation by developers and designers.

Another interesting result that should not pass unnoticed is the importance of including blind users during the evaluation sessions of websites. As the two user groups encounter largely different problems, addressing the problems sighted users have would not necessarily mean the problems blind users have will be addressed.

10.2. Research contributions

This section presents what was done to answer each research question asked at the beginning of the thesis as well as the research contributions.

10.2.1. What are the most frequent components that researchers consider as part of the concept of web accessibility?

The first contribution of this thesis is theoretical. It provides a better understanding of what researchers consider to be the key components of the concept of web accessibility.

This was achieved by conducting an analytical study that draws together the research literature into a common unified definition of web accessibility. The unified definition of web accessibility proposed encompasses the most frequent components considered by researchers. Based on the analysis, the following unified definition of web accessibility was proposed:

"all people, particularly disabled and older people, can use websites in a range of contexts of use, including mainstream and assistive technologies; to achieve this, websites need to be designed and developed to support usability across these contexts".

The unified definition can be used to provide clarity for what researchers control in their studies and in terms of what they measure when they investigate the concept. In addition, it can help the researchers making studies more comparable. The unified definition was used as the basis of the next studies of this thesis that operationalise the concept.

10.2.2. Which verbal protocol can be considered a better option for user-based studies with blind and sighted users on the web?

The second contribution of this thesis is methodological as it provides a better understanding of which verbal protocol can be considered a better option for eliciting problems on the web for both blind and sighted users.

A study with 16 participants, eight blind and eight sighted, was conducted by performing an evaluation of the two verbal protocols (CVP and RVP). The results of the study indicated that RVP could be considered a better option for eliciting problems on the web for both blind and sighted users. RVP identified more distinct problems but also more interactivity type problems than CVP, a problem type that was the most frequently reported by either user group. However, RVP comes with a drawback as it found that it demanded greater workload for participants than CVP.

This study was the first, to our knowledge, that compared the two verbal protocols with blind users. The study provided additional knowledge in regard to which verbal protocol can be considered a better option for identifying problems on the web. The study findings in relation to the differences of the number of problems identified between the two protocols are in line with the findings of previous studies (Van den Haak et al., 2003, 2004), however, these studies were conducted only with sighted users. The study findings in regards to the overlapping figure of the problems between the two verbal protocols was similar with the figure found in previous studies with sighted users (Van den Haak et al., 2007, 2009).

The results of the study in Chapter 4 expand our knowledge of which verbal protocol can be considered a better option in studies with blind and sighted users. Based on the results, RVP can be considered a better option in user-based studies, particularly if the interest is in interactivity problems. However, for studies interested in content or information architecture problems, either protocol can be considered appropriate. The result of this study guided the verbal protocol used in the next study of the thesis, which was RVP. The focus of the study was to explore the problems between blind and sighted users in websites and this included the spectrum of problems users may encounter, such as interactivity problems.

10.2.3. What are the problem similarities and differences between blind and sighted users on the web?

The third contribution of this thesis is an empirical evidence-based research of the type of problems that are distinct and shared between blind and sighted users on the web. Although there is an indication in the literature of known overlap and differences of problems between blind and sighted users, there is little knowledge of what these problems are and what causes them. This study is the first, to our knowledge, which conducted a comparison of the problem differences between the two user groups at this level.

The problems found on the web from 24 participants, 12 blind and 12 sighted, were analysed to investigate their problem similarities and differences between the two user groups. The work contributes to the existing knowledge of the problems between blind and sighted users on the web by corroborating the results of Petrie and Kheir (2007) and expanding our knowledge of the problem differences between the two user groups.

Two key problem types that were also distinct to blind users were the poor page structure and the lack of feedback on users' actions. Both problems can play an important role in users' experience and are probably of the main sources of their low task success rates. Other problems included the excessive effort required by blind users due to the refreshing of the page on users' actions, interactive elements not labelled and not properly implemented using the appropriate markup, headings not being descriptive, overwhelming content areas and screen reader incompatibilities with the web page content. An important finding of this study is that the two user groups may encounter similar types of problems. However, the problem characteristics largely differ, with many problem differences have to do with users' navigation differences. Another interesting finding was that many of the key problems blind users encounter were around a specific design aspect, the search and filtering browsing of content. Although this is a very specific design aspect, it is in a substantial number of websites on the web. Problems related to this specific design aspect seems to be on other shopping websites as well, as the websites used in this study had very similar design features with many shopping websites on the web.

This work provides a few important implications for website designers and developers. First, it expands our knowledge of how problems are distinct between blind and sighted users and provide a better understanding of the range and diversity of user problems. The large problems difference also highlights the importance of including blind users during evaluation sessions of websites, as addressing only the problems sighted users on the web will not necessary means the problems blind users have will be addressed. Moreover, it shows that there are gaps in existing website designs, as websites do not accommodate the needs of blind users and do not provide all the necessary design features to create an accessible user experience. The study pointed out a few key problems that worth further investigation, particularly the problems related to the search and filtering browsing of page content.

10.2.4. What are the benefits of specific design solutions to the key problems on blind users' experience?

Another empirical evidence-based contribution of this thesis work is the benefits of blind users' experience on the web by addressing the key problems they encounter through specific design solutions. These were achieved by conducting two studies (Chapter 8 and Chapter 9), with 18 and 20 blind users respectively. The first study (Chapter 8) was conducted in a non-exploratory browse and search environment. The second study was a confirmative study of the findings of the first study (Chapter 9) with participants doing an exploratory ecological task.

Based on the results, the page structure is of high importance in order to improve blind users' effectiveness on the web. Participants were able to perform the tasks with very high success rates, with more than 90% on all website designs, in both studies. This result highlights the importance of the page structure and following the accessibility standards for organising the page content. Furthermore, addressing the problems with the lack of feedback on users' actions only made it clearer about what is happening on the page in response to users' actions. This is because the constant refreshing of the page each time users select a filtering option seems to overshadow all the benefits the extra feedback provides. Addressing the problems related to the excessive effort required due to the refreshing of the page and the lack of feedback on users' actions can contribute to an improvement of the overall user experience, increases users' efficiency, is easier to use, and reduces the overall workload as well as specific workload measures.

These two studies provided a better understanding of how common interactivity problems influence the effectiveness, efficiency and the perceived usability of a website for blind users. Also, they provide a better understanding of how specific design solutions can address these interactivity problems to improve the qualities of usability on a website. Moreover, it shows that simple design solutions may address some of the problems encountered by blind users, but may not be enough to improve the overall users' experience. However, a combination of several design solutions can provide a cumulative improvement in experience. This last result shows that solving these problems is quite complex as they influence one another.

10.3. Limitations across studies

There are a few limitations in this thesis that need to be considered. All empirical studies conducted were primarily focused on users' experience measures. The studies did not engage with other measures, such as participants' keystrokes. Keystrokes can be very informative, particularly when conducting studies with blind users, as they can provide insights into users' navigation techniques. For example, what screen reader keys users prefer to use for navigating through the page content. However, when eliciting user problems on the web, keystrokes are rarely collected (Coyne & Nielsen, 2001; Disability Rights Commission, 2004; André Pimenta Freire, 2012; Lazar et al., 2012; Petrie & Kheir, 2007). Analysing users' keystrokes can be a very time-consuming process and evaluators should carefully consider whether they actually need this data and plan ahead for the additional time they will need to analyse the data (Dumas & Redish, 1999; Rubin, 1994). When evaluators look into the problems users have on a website, they rarely need users' keystrokes to understand users' problems (Dumas & Redish, 1999). Thus, users' keystrokes were not collected in the studies of this thesis. However, keystrokes would have been beneficial to provide insights into users' navigation techniques, particularly when different website designs are evaluated, as it would have been interesting to know whether their navigation techniques were affected between the different website designs.

Another limitation of the studies is that participants did not use their own equipment, which threats the ecological validity. There were some important considerations that did not allow this choice. First, I would not be able to assess whether the equipment was in running order before the arrival of the participant. Second, the sessions in all studies were recorded using software that was preinstalled on the computers used in the studies. Third, it would have limited the participant pool to ones that had a laptop computer at home or at work. Fourth, it would have possibly made participants' commuting less comfortable and less convenient. For these reasons, it was preferred for participants not use their own equipment.

Although this thesis work provided a better understanding of the problem differences between blind and sighted users on the web, there is an important limitation that needs to be considered. The websites that were used in the study in Chapter 5, can be classified as shopping websites. Although, it is not disagreed that some problems can occur in other type of websites, there are some problems that may not be able to be generalised to other type of websites. For example, the lack of feedback on users' actions when they interact with the filtering options of the page, is a problem that would be explicit to shopping type websites. However, that does not disregard that blind users might have had problems with the lack of feedback when they interact with the page content on other types of websites.

All studies were conducted using desktop websites rather than using mobile devices and that may threaten the external validity, as some of the problems may not apply on mobile devices. It was preferred not to use mobile devices to avoid any conflicts between the presentation of the website on different screen sizes. For this reason, it was preferred to conduct the studies on desktop websites, where the presentation of the website content was going to be consistent.

Another limitation is that the number of websites and tasks used to elicit users' problems in studies Chapter 4 and Chapter 5 was low. A larger number of websites was not preferred as each website evaluation with blind users took around an hour to conduct. Therefore, to mitigate any fatigue effects, the number of websites selected in the studies was low.

The most important limitation in all empirical studies is that some of the participants took part in multiple studies, which can threaten the internal validity of the studies. This was a pragmatic limitation due to the difficulties of recruiting participants from a small participant pool. However, this is a learning point on how to conduct research in this field area. Different approaches can be taken, in order to mitigate any impacts of recruiting participants in multiple studies. For example, researchers need to have a considerable time gap between the studies undertaken by the same participants. When recruiting participants, the researcher can have screening questions, such as when was the last time a participant took part in a research study and of what nature. Look into differences when the researcher analyses the data, to check whether there are any impacts due to the participation in another research. Keep an up-to-date database with all the participants that includes information about the studies participants took part. The researcher can first reach to participants that were not involved in a similar type of studies or were not involved in a study recently. All these considerations are valuable lessons from the whole thesis on how to conduct studies when recruiting from a small participants pool. Moreover, to make the participation easier, the participants should be compensated for their time and reimbursed for their travel expenses. The researcher should also assist participants with their travel if need be.

10.4. Future work

While the research in this thesis has provided insights into the problems blind users have on the web and tested specific designs solutions to some of the prevalent problems, there is abundant room for further progress in improving blind users' experience on the web.

Future research could be done to investigate the problems blind users had with the page content. Users encountered many problems with irrelevant content or too much content. First, it is unclear what content can be considered irrelevant to users, as users many times will assess the relevance of the content to their task. Advertisements will probably be considered irrelevant content, but there were also other sections on the page that users found the page having irrelevant content. Moreover, it is not clear whether users will still perceive this content as relevant or not if the pages are well structured. As participants workload will be higher in pages with poor structure, the problems with the content they encountered may be because of the high cognitive load the website and task imposed on users. It is unclear whether the content itself was reducing the users' experience. Future research can be done to test whether users still perceive content as irrelevant to their task in websites with good structure, and if so to investigate what content is considered relevant or not.

Another area worth further investigation revolves around the issues with the order that blind users perceive the content on the page. Like the previous problem, it is not clear whether the order that the content was positioned on the page was the problem or the poor page structure made it difficult for users to find it on the page. In addition, blind users navigate in one dimension (i.e. top to bottom) which can cause information to be perceived in a different order in comparison to sighted users who navigate in two dimensions (i.e. top to bottom, left to right in English language pages). Moreover, these issues may have to do with how blind users expected the websites to be structured. Future work worth exploring this issue in order to provide a better understanding of these problems as well as how blind users expect the page content to be structured.

Many of the problems blind users encountered were imposing high cognitive workload on users. However, to reduce users' workload it was found that a combination of design solutions is necessary as simple design solutions are not sufficient. This result leaves room for further research for researchers and designers to identify other areas in websites that impose high cognitive load to blind users and propose and test design solutions that mitigate it.

While this thesis provided a further understanding of the problems blind users had on the web and tested specific design solutions to some of the key problems they encounter, it is important to conduct further research into design solutions that are accompanied with empirical evidence of their benefits on users' experience.

10.5. Concluding remarks

The research presented in this thesis contributes to the area of accessibility. The research contribution is threefold: empirical, theoretical and methodological. It provides empirical insights into the problem similarities and differences between blind and sighted users on the web. Also, it tests specific design solutions to solve key problems blind users have and contributes to an understanding of their benefits on users' experience. This work also provides a theoretical contribution to the field of web accessibility by proposing a unified definition of web accessibility that can be used in future research that operationalise the concept. It also provides a methodological contribution to which verbal protocol can be considered a better option in user-based studies for eliciting problems on the web with blind users.

Appendices

Appendix A. Studies looking into the problems blind users have on the web

Table 56. Studies that look into the problems that blind users have on the web.

Reference	No. websites	Websites Domain	No. blind participants	Other user groups (N)	Sighted users	Verbal Protocol	Definition	Reference to Standards
Oppenheim and Selby (1999)	3	search engines	1	partially sighted (3)	-	N/S	N/S	W3C, Section 508
Coyne and Nielsen (2001)	16	Government, e-commerce, business, non-profit sites	18	partially sighted (17), physical impaired (9)	-	CVP	N/S	W3C, Section 508
	4	Unspecified, e-commerce, information	20	partially sighted (20)	20	CVP	N/S	W3C, Section 508

Byerley and Beth Chambers (2002)	2	Libraries	2	-	-	CVP	N/S	W3C, Section 508
Craven (2003)	4	Search engine, library, e-commerce, information	10	partially sighted (10)	20	N/S	N/S	W3C
Disability Rights Commission (2004)	100	General	10	partially sighted (10), dyslexic (10), physical impaired (10), hearing	-	N/S	N/S	W3C, Section 508

				impaired (10)				
	6	General	10	-	10	N/S	N/S	W3C, Section 508
Federici et al. (2005)*	1	Education	N/S	visual disability (4), physical impaired (2)	-	CVP	N/S	W3C
Brebner and Parkinson (2006)	6	Library	3	partially sighted (6)	-	N/S	Own definition	W3C, Section 508
Jaeger (2006)**	10	Government	N/S	"ten participants had either visual	-	CVP	N/S	W3C, Section 508

				impairment or mobility impairment"				
Petrie and Kheir (2007)	2	Mobile	6	-	6	CVP	Own definition	W3C, ISO 9241
Rømen and Svanæs (2008, 2012)	2	Government	3	physical impaired (2), dyslexic (2)	6	CVP	N/S	W3C
Stenitzer et al. (2008)**	5	e-commerce	N/S	older people, partially sighted	-	N/S	N/S	-
R. Babu and Singh (2009)	1	Virtual Learning Environment	6	-	-	CVP	N/S	W3C, ISO

Power, Petrie, Sakharov, and Swallow (2010)	3	Virtual Learning Environment	4	-	-	CVP	N/S	W3C
Giraud et al. (2011)	2	Social, e-commerce	6	-	6	N/S	N/S	W3C
Swierenga et al. (2011)	1	Information	8	partially sighted (8)	18	N/S	N/S	ISO 9241
André Pimenta Freire (2012); Power et al. (2012)	16	General	32	partially sighted (19), dyslexic (13)	-	CVP	ISO 9241	W3C, Section 508, ISO 9241
Lazar et al. (2012)	16	Job seeking	16	-	-	CVP	N/S	W3C, Section 508

Abu-Doush et al. (2013)	8	Government	20	-	-	CVP	Abanumy et al. (2005)	W3C, Section 508
R. Babu (2013)	1	Travel	5	-	-	CVP	N/S	-
R. Babu and Singh (2013a)	1	Social	5	-	-	CVP	N/S	W3C
R. Babu and Singh (2013b)	1	Virtual Learning Environment	6	-	-	CVP	N/S	W3C, Section 508
Ramayah et al. (2013)	1	Social	1	-	-	N/S	N/S	W3C, Section 508
Yoon et al. (2013)	5	Library, general	6	-	N/S	CVP	N/S	W3C, Section 508

Coursaris et al. (2014)	1	Health	16	partially sighted (4)	5	CVP	ISO 9241	W3C, Section 508, ISO 9241

* The number of blind participants is not clear, as the study does not provide adequate information

** The number of blind participants has not been reported

Table 57. Problems found from the studies that look into the problems blind users have on the web.

Reference	Problems found
Oppenheim and Selby (1999)	• interpretation of speech synthesis was sometimes difficult to understand • links not descriptive • repeated information • irrelevant content • content not descriptive • images without alternative text • images with non-descriptive alternative text • page structure not clear
Coyne and Nielsen (2001)	• images without alternative text • images with non-descriptive alternative text • buttons description not clear • pop-ups not accessible • difficulties skipping content on the page • links open at new browser windows without any indication • too many links • instructions using sensory characteristics • not easy to navigate through the page • irrelevant content • error messages conveyed through colour only • required form fields not clear • content not in appropriate order

	• content not descriptive • use of tables for visual design instead to organise in- formation
Byerley and Beth Chambers (2002)	• links not descriptive • content not well structured • table not well structured • no skip link • images with non-descriptive alternative text
Craven (2003)*	• too much information • duplicate information • content not descriptive • structure not clear • complicated page structure
Disability Rights Commission (2004)	• screen reader incompatibilities with the page content • links not descriptive • no labels associated with input controls • no titles on frames • complicated page structure • images without alternative text • images with non-descriptive alternative text
Federici et al. (2005)	• links not descriptive • too many links • content available only in PDF format • no skip link • content confusing

Brebner and Parkinson (2006)	• links not descriptive • no skip link • no labels associated with input controls • images with non-descriptive alternative text
Jaeger (2006)	problems reported are not organised per user group
Petrie and Kheir (2007)	does not provide details about the problems found
Rømen and Svanæs (2008, 2012)	• similar links • links not descriptive • too many links • redundant links • pop-ups not accessible • lack of instructions on how to use the forms
Stenitzer et al. (2008)**	• missing or inadequate labels for links and buttons • position of elements not following users' expectations • important information not positioned at the top of the page • too much information on the page • navigation through the page content was difficult • disturbing advertisements
R. Babu and Singh (2009)	• lack of feedback from screen reader when users arrive into a new page

	• functionality not as expected. Screen reader focus was not at the top of the page when the page refreshed.
Power et al. (2010)	provides only the number of problems blind users encounter
Giraud et al. (2011)	• content not clear • irrelevant content • content not well structured • no feedback on user's actions (activated links)
Swierenga et al. (2011)	provides design recommendations for each user group, however it is not clear whether the recommendations derive directly from issues that users encounter on the websites
André Pimenta Freire (2012); Power et al. (2012)	• links not descriptive • navigation elements do not help users find what they are looking for • content not found where expected by users • irrelevant content • controls not clear what will do • no labels associated with input controls • no feedback to inform an action has had an effect • content not clear • functionality not working • areas inaccessible via screen reader • no headings

	• images with non-descriptive alternative text • no enhancement to audio, video, or multimedia
Lazar et al. (2012)	• links not descriptive • complicated page structure • screen reader incompatibilities with the page content • required form fields not clear • areas inaccessible via screen reader • functionality missing • illogical heading structure • data input not clear • table not well structured • no labels associated with input controls • labels/instructions not clear • error messages do not help user recover from their errors • content not in appropriate order
Abu-Doush et al. (2013)	• areas inaccessible via screen reader • images without alternative text • links not descriptive • links open at a new tab without any indication • not easy to navigate through the page • too many links
R. Babu (2013)	• labels/instructions not clear • no feedback on user’s action with buttons

	• no labels associated with input controls
R. Babu and Singh (2013b)	• no feedback on user's actions with links • feedback on user actions is confusing • pop-ups not accessible • links not descriptive • content not well structured
R. Babu and Singh (2013a)	• labels/instructions not clear • lack of instructions on input entries • error messages do not help user recover from their errors
Ramayah et al. (2013)	• content not well structured • complicated page structure
Yoon et al. (2013)	• too much information • complicated page structure
Coursaris et al. (2014)	• content not clear • pop-ups not accessible • no feedback on user's actions • functionality not as expected • abbreviations not explained • difficulties interacting with input controls • no labels associated with input controls

* The problems reported are from both blind and partially sighted users.

** The problems reported are from older people, partially sighted and blind users.

Appendix B. Web Accessibility Definitions List

The definitions are organised in two groups in chronological order. Definitions from standards and definitions from books, papers and online documents.

Standards

[1] Technology is accessible if it can be used as effectively by people with disabilities as by those without. (Section 508, 1996)

[2] ... people with disabilities can use the Web. More specifically, Web accessibility means that people with disabilities can perceive, understand, navigate, and interact with the Web, and that they can contribute to the Web. (WAI, 2005)

[3] ... the usability of a product, service, environment or facility by people with the widest range of capabilities. (ISO 9241-171, 2008)

[4] usability of a product, service, environment or facility by people within the widest range of capabilities. The concept of accessibility addresses the full range of user capabilities and is not limited to users who are formally recognized as having disability. The usability-oriented concept of accessibility aims to achieve levels of effectiveness, efficiency and satisfaction that are as high as possible considering the specified context of use, while paying attention to the full range of capabilities within the user population. In a web context, accessibility means the degree to which people with disabilities can perceive, understand,

navigate, and interact with the web, and that they can contribute to the web. (BSI, 2010)

Books, papers, online documents

[5] ... the ability for [web] browsers to render information in a manner that is accessible to people with disabilities. For the blind, any aspect of a graphic interface presents barriers. For low vision web surfers (and in some cases, those with cognitive limitations), data presentation in different formats, different fonts, and inconsistent character and word spacing, make reading online information difficult. For the deaf, rendering sounds or sound bytes presents significant challenges. (Paciello, 1996c)

[6] .. it is critical that the Web be usable by anyone, regardless of individual capabilities and disabilities. (Berners-Lee, 1997)

[7] ... anyone using any kind of web browsing technology must be able to visit any site and get a full and complete understanding of the information as well as have the full and complete ability to interact with the site if that is necessary. (Letourneau, 1998)

[8] ...the design of a webpage ... in order to ensure that all users can access the information on the page. (Waddell, 1998)

[9] Web accessibility is the ability for a person using any user agent (software or hardware that retrieves and renders web content) to understand and fully interact with a website’s content. The idea of accessibility is based on more than the implementation of standards; it

embodies the idea that everyone has the right to be included in society, regardless of disability, geographical location, language barriers, or any other factor. (Sierkowski, 2002)

[10] ... individuals with disabilities can access and use them [websites] as effectively as people who don't have disabilities. (Slatin & Rush, 2002)

[11] ... people being able to get and use web content. It is about designing web pages that people can present and interact with according to their needs and preferences. A primary focus of accessibility is access by people with disabilities. The larger scope of accessibility includes benefits to people without disabilities ... Accessibility is a subset of a more general pursuit: usability. (Thatcher et al., 2002)

[12] Web Accessibility refers to the possibility of accessing any web content by anyone regardless to circumstances such as impairments, platforms, devices, browsers, etc. (Abascal, Arrue, Garay, & Tomás, 2003)

[13] An accessible web site is a web site that can be successfully used by people with various disabilities. People with different disabilities may be using different forms of assistive technology, such as screen readers, alternative keyboards, or alternative pointing devices. A web site that is accessible is flexible enough to work with these various assistive technology devices. (Lazar, Schroeder-Thomas, et al., 2003)

[14] The bottom line with respect to web accessibility is whether an individual can perform a website's intended function(s). As there will be varying degrees in the ease with which users can do so, such a measure does

not lend itself to a binary “approved” or “not approved” rating. With this in mind, the evaluator of any web page should (a) identify its perceived intended function(s) and (b) rate the page on a scale that measures the ease with which any user, including a user with a disability, can perform the intended function(s). (Thompson, Burgstahler, & Comden, 2003)

[15] An accessible Web site is one that can be used by people with disabilities. People with disabilities may use assistive technologies such as screen readers, Braille printers, and alternative pointing devices. In addition, they may also adjust graphical browsers to improve accessibility; however, this is only effective if the Website is designed to be flexible and accessible. To be accessible, a Website must be flexible enough to work with the various assistive technology devices that a person with a disability might use and to provide the relevant content in an accessible modality. (Lazar, Beere, et al., 2003)

[16] Accessibility, when pertaining to a Web page, means that information has been made available for use by almost everyone, including persons. with disabilities. This accessibility may be direct or through the use of assistive technologies. (Hackett, Parmanto, & Zeng, 2004)

[17] In its broadest definition, "web accessibility" is an approach to web design that aims for maximal inclusion, both in terms of people who use web sites, and the technologies that are utilised in the process….These days Web accessibility generally refers to accessibility for disabled user groups. (Alexander, 2004)

[18] Web accessibility can be defined simply as to which degree a site is accessible to the largest possible range of people. The more people are able to access a website, the more accessible is the site. At its core, Web accessibility emphasizes making website accessible to persons with disabilities and involves removing potential barriers to access caused by inconsiderate website designs......Web accessibility can be defined as the degree to which it is accessible through assistive technologies used by persons with disabilities. (Zeng, 2004)

[19] Web accessibility refers to the degree to which web information is accessible to all human beings and automatic tools. The goal of web accessibility is to allow universal access to information on the web, by all people but especially by people with any impairment, no matter what its severity, (e.g. blindness, low vision, deafness, hard of hearing, physical disabilities or cognitive disabilities). In addition, the information must be accessible by automatic machine tools. (Abanumy, Al-Badi, & Mayhew, 2005)

[20] Web accessibility involves making web content available to all individuals, regardless of any disabilities or environmental constraints they experience. (Mankoff et al., 2005)

[21] ... the most important component in web accessibility is addressing issues relevant to individuals with disabilities and the elderly. (Maswera et al., 2005)

[22] ... in general, accessibility can be defined as the ability of anyone, including those who have disabilities, to access content and information on the Internet. (Mancini, Zedda, & Barbaro, 2005)

[23] Web accessibility measures how easily diverse sets of users, regardless of disability or environmental constraints, can access material on a website. (Bailey & Burd, 2005)

[24] Accessibility here means that people with disabilities are offered the opportunity to access web content easily. In order to achieve this, information must be available for different devices and platforms. This means that the coding has to follow some basic rules, and that any information must be accompanied by metadata. (Zerfass & Hartmann, 2005)

[25] ... accessibility is defined as the practice of giving people with disabilities equal access to websites and online content. (Holsapple, Pakath, & Sasidharan, 2005)

[26] Accessibility aims to allow the access to the content of the Web application even in presence of reduced hardware/software configurations on the client side of the application (such as browser configurations disabling graphical visualization, or scripting execution), or of users with physical disabilities (such as blind people). (Di Lucca, Fasolino, & Tramontana, 2005)

[27] “Accessibility” refers to the extent by which the web site, including the technology such as hypertext coding, is barrier-free to all users of the

information, thus providing enhancements that enable people with disabilities to move towards independence....web accessibility is the ability for a person using any agent (software or hardware that retrieves and renders web content) to understand and fully interact with a web site's content. (Yates, 2005)

[28] Sometimes accessibility is defined in terms of effectiveness; now and then it is defined in terms of usability; but unfortunately there are too often claims that a web site is accessible simply because an automatic accessibility testing tool yielded no error. (Brajnik, 2006)

[29] In general, accessible websites are able to give everyone equal opportunities to access the complete Web content regardless of software, hardware and user ability. (Chen, Chen, & Shao, 2006)

[30] ... Web accessibility means that people with disabilities can easily navigate and interact with the Web. (Iaccarino, Malandrino, & Scarano, 2006)

[31] For a web site to be accessible, it should provide equal or equivalent access to all users, and it should work compatibly with assistive technologies such as narrators, screen enlargement, and many other devices that persons with disabilities may employ to navigate cyberspace. (Jaeger, 2006)

[32] .. refers to the extent that all users are able to successfully gain access to information presented. (Brebner & Parkinson, 2006)

[33] The term web accessibility generally refers to the ability of people to access the World Wide Web. The application of technical solutions to the design of a website is good practice which aims to improve accessibility - particularly for people who use assistive technologies, such as screen readers, screen magnification, or electronic Braille. (Craven, 2006)

[34] Accessible Web pages accommodate the differing capabilities, needs and situational considerations of Web users…. Specifically, accommodating the needs and capabilities of the disabled means that Webpages must be designed to allow the effective use of assistive technologies. (Williams, Rattray, & Grimes, 2006)

[35] ..the affordance of an interactive application to be used effectively and efficiently by people with special needs. (Bolchini, Colazzo, & Paolini, 2006)

[36] The accessibility of websites on the Internet can be defined as the combination of technologies and norms for implementing and designing them, which facilitate the use of the Internet for the largest possible number of people, including those with disabilities. In this last case, various types of disabilities which can hinder or prevent the person from having access to the information and making effective use of the website, have been identified. These not only include visual, auditory and ones related to movement, but also learning disabilities (this includes many older people who are technologically illiterate), and 'technological handicap' due to a lack of adequate technical means to gain access to all areas of the website (equipment which is not up-to-

date regarding the latest hardware/software technologies or lines with slow access). (Joaquín Mira, Llinás, Tomás, & Pérez-Jover, 2006)

[37] Accessibility is concerned with making information on websites available to the widest audience possible; while this includes users with disabilities, application of accessible design principles should improve the online experience of all users. (Paris, 2006)

[38] The term web accessibility can refer to the provision of physical access to appropriate hardware and software to enable access to the web; it can mean the provision of add-on technologies to widen access to the web, for example through the use of assistive technologies such as screen reading software, screen magnification, alternative mouse devices, and voice input. Web accessibility can also refer to the design of the web interface which, according to recommended standards and guidelines, should be presented in a way that can be interpreted by as wide a group of user as possible and by any kind of assistive technology. (Craven & Nietzio, 2007)

[39] .. means that people with disabilities can use a product. (Henry, 2007)

[40] Web accessibility refers to the degree to which a website may be accessed by people with varying abilities. (S. K. Kane, 2007)

[41] ... the ultimate criteria for accessibility should be user-based and we can adapt the ISO 9241 definition for this purpose: the extent to which a product/website can be used by specified users with specified

disabilities to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use. (Petrie & Kheir, 2007)

[42] Web accessibility means the ability [of websites] to be accessed by all kinds of people or devices. (M. Watanabe et al., 2007)

[43] Accessibility is a subset of criteria from a wider purpose: usability... Accessibility is related to make a system usable, efficient, effective and to satisfy “more people in more different situations”. In other words, making Web more accessible means to concern about providing content access to people with different abilities, using different devices, including assistive technologies, such as screen readers, non-conventional input devices and others. (De Oliveira Junior et al., 2007)

[44] Web accessibility can be defined simply as to which degree a site is accessible to the largest possible range of people. The more people able to access a website, the more accessible is the site. At its core, web accessibility emphasizes making websites accessible to persons with disabilities and involves removing potential barriers to access caused by inconsiderate website designs. (Curran, Walters, & Robinson, 2007)

[45] In general, accessibility means the ability to access. As for an accessible website, the website must allow an equal opportunity to everyone in order to access all available content regardless of software, hardware and user ability. (Mitsamarn, Gestubtim, & Junnatas, 2007)

[46] The terms ‘accessibility’ and ‘usability’ have distinct, but related, meanings. Web accessibility is the measure of usability when users are constrained by any of a number of disabilities: • Physical accessibility barriers, including blindness and limited mobility • Cognitive barriers, relating to the brain and mental processes or • Circumstantial barriers, relating to the kinds of devices used to access the web. These could be devices with limited screen size, memory, or bandwidth, such as mobile phones or PDAs (Adapted from: W3C, 2006) So, accessibility is partly about providing the same information to all users, but in a number of different ways. (D. Kane & Hegarty, 2007)

[47] Accessibility means access to information for all – focusing on people with disabilities and senior citizens. Ensuring accessibility improves the quality of life for such people by removing barriers that prevent them from taking part in many important life activities. (Suh & CHO, 2007)

[48] Web accessibility is the practice of making Web sites accessible to people who require more than just traditional Web browsers to access the Internet. For example, a visually impaired user can use a screen reader to translate text and graphics on the computer screen to an audio format so the user hears the screen content via a speech synthesizer or sound card. An accessible Web site is designed to accommodate a wider set of ways users can access the site. (Peters & Bradbard, 2007)

[49] Web accessibility means that people with visual, physical, speech, cognitive or neurological disabilities are given the opportunity to perceive, understand, navigate and interact with the Web. (De Lima, Lima, & De Oliveira, 2007)

[50] ...the goal of web accessibility is to allow universal access to information on the web, by all people but especially by people with any impairment, no matter what its severity. (Lee, Kim, & Kim, 2007)

[51] ... Web accessibility is important to ensure interoperability between different applications and to enable users to access the Web using their preferred format. This could be via assistive technology to interact directly with the site or to download information into an alternative format....In the literature Web accessibility generally refers to the application of technical solutions to the design of a Web site in order to render it more accessible to users, in particular users of assistive technologies. (Brophy & Craven, 2007)

[52] ... to increase usability for people with disabilities and in scenarios involving mobile and embedded devices. (Wendy Chisholm & May, 2008)

[53] Web accessibility aims to help these people [who have disabilities] to perceive, understand, navigate, and interact with, as well as contribute to, the Web, and thereby the society in general. This accessibility is, in part, facilitated by the Web Content Accessibility Guidelines (WCAG) currently moving from version one to two. (Harper & Yesilada, 2008)

[54] Accessibility is a concept related to providing access to Web content to people with different abilities and people using different devices. (André P Freire, Fortes, Turine, & Paiva, 2008)

[55] ... it is important that the information be easily reachable by all, including people with disabilities... Accessibility is aimed specifically at making Web sites more available to a wider population of users (including special categories) by removing the technical barriers that prevent access to the information included in the site. (Leporini & Paternò, 2008)

[56] An accessible website is defined as one that ensures that all of its pages can be used effectively by all persons using that website. (Mills, Han, & Clay, 2008)

[57] Web accessibility has become an important issue in Web development. Making Web more accessible is related to providing access to content to people with different abilities, using different devices, including assistive technologies, such as screen readers, non-conventional input devices and others. (André P Freire, Bittar, & Fortes, 2008)

[58] ... is one that is sufficiently flexible to be used by all people including those using assistive technologies such as; screen readers, voice browsers and Braille displays. (Baguma & Lubega, 2008)

[59] The concept of Web accessibility is related to the possibility to enable any user, using any user agent (software or hardware to display Web content) to understand and fully interact with a Web site, despite disabilities, languages or technological constraints. (A. P. Freire, C. M. Russo, & R. Fortes, 2008)

[60] Accessibility is the equal access to information and communication technologies (ICTs) for individuals with disabilities, and it is of utmost

importance to persons with disabilities in the networked society. Accessibility allows individuals with disabilities, regardless of the types of disabilities they have, to use ICTs, such as Websites, in a manner that is equal to the use enjoyed by others. (Jaeger, 2008)

[61] A website is said to be accessible when anyone, regardless of economic, geographic or physical circumstances, is able to access it. (Good, 2008)

[62] Website accessibility is defined as the ability of the website to serve user with disabilities, especially blind people in accessing the internet....website accessibility is the ability of the website to be accessed by user using all of the existing browser technology. The website accessibility also refers to the capability of user to understand all of the information contained in the website and the ability of user to interact with the website if it is needed. (Jati & Dominic, 2008)

[63] ...Web accessibility is about ensuring that anyone, using any browser or device is able to access any content on the Web. This definition is consistent with Letourneau's (1998) position that accessibility ought to be concerned with ensuring that all users (regardless of ability) should be able to access sites using current and legacy browsers as well as emerging non-browser technologies, and gain full and complete understanding of the content of those sites. (Friedel & Wood, 2008)

[64] ...Web accessibility corresponds to making possible to any user, using any user agent (software or hardware to view Web content) to

understand and interact with a Web site, despite of disabilities, languages or technological constraints. (Andre P Freire et al., 2008)

[65] ...website is accessible when specific users with specific disabilities can use it to achieve specific goals with the same effectiveness, safety and security as non-disabled people. (Brajnik, 2008)

[66] An accessible Web site is a site that can be perceived, operated, and understood by individual users despite their congenital or induced disabilities ... It means having a web application usable to a wide range of people with disabilities, including blindness and low vision, deafness and hearing loss, learning difficulties, cognitive limitations, limited movement, speech difficulties, photo-sensitivity and combination of these. In short, we can say that accessibility addresses a universal usability. (Martín, Cechich, & Rossi, 2008)

[67] Web accessibility can be defined as a person's ability to access the Web. A Web site is accessible if it can be used as effectively by people with disabilities as well as by other people. The contents of the site, the facilities and services should be accessible to a wider audience as possible, regardless of age, disability or limitations of the technology or the environment of the end user.....Accessibility means in broad sense Web for all. Accessibility in the strict sense means taking into consideration of people with disabilities. (Jitaru & Alexandru, 2008)

[68] ... Web accessibility can be defined as making Web resources accessible to all users, regardless of the technical, physical or mental restrictions on the client side. This means that accessible Web sites are

aimed at being equally accessible for all people – disabled or not. It also implies that a Web site has to work irrespectively of the browsing technology on the client side. Furthermore it is stated that a Web site has to be inclusive and attractive to users and has also to make its benefits visible to potential users. (Kern, 2008)

[69] Web accessibility means that people with disabilities, including older people with changing abilities, can perceive, understand, navigate, and interact with the Web. (Borrino, Furini, & Roccetti, 2009)

[70] The main goal of Web Accessibility is to make it possible for everyone to use, understand and communicate using Web based resources, despite any disabilities or technological restrictions. (Batra, 2009)

[71] Web accessibility is the degree to which people with visual, auditory, physical, speech, cognitive, or neurological disabilities can perceive, understand, navigate, and interact with the Web. (De Lima et al., 2009)

[72] The term accessibility, as applied to the Internet, means that anyone can equally access the information presented, regardless of device and/or personal limitations. (Hackett & Parmanto, 2009)

[73] The aim of Web accessibility is to guarantee that Web applications can be acceded and used by all potential users independently of the limitations of the individuals themselves or the derivatives of the context of use. Therefore, it includes the use for persons with all kinds of physical, sensory, and cognitive characteristics, with any browser (current, old), with any kind of computer, with any type of connection,

and with any type of device (TV, mobile phone, etc.). (Andrés, Lorca, & Martínez, 2009)

[74] An accessible website is one that is sufficiently flexible to be used by all people including People with Disabilities (PWDs). Although accessibility is a vital quality attribute for PWDs, it has not yet gained much recognition as a crucial non-functional requirement like security, performance, accuracy and usability. (Baguma, Stone, Lubega, & van der Weide, 2009)

[75] The term accessibility can easily be defined has the possibility of disabled people interact with a product, resource, service or activity has normal people would. In what concerns the ICT, we can define accessibility as the creation of interfaces that are perceived, operable and easy to understand for people with a wide range of features. This includes all deficiencies, functional limitations, including a visual impairment, hearing, physical, cognitive and neurological. In this set should also be included conditions of temporary incapacity, such as the loss of glasses or the breaking of an arm. Beyond this, accessibility also makes the products more accessible to people who do not have any kind of disability. (Martins, Cruz, & Gonçalves, 2009)

[76] Accessibility of user interfaces can be approached through a usability field.... Accessibility focuses on including people with disabilities as the "specified users" and a wide range of situations, including assistive technologies (ATs), as the "specified context of use". In a simpler way, usability means designing a user interface that is effective, efficient, and satisfying. Accessibility makes sure the user interface is designed to be

effective, efficient, and satisfying for more people—especially people with disabilities, in more situations—including with ATs. (Moreno, Martínez, & Ruiz-Mezcua, 2009)

[77] Website accessibility can be defined as the ability to access the web regardless of "visual, hearing, mobility or learning disabilities", speed of Internet connection/bandwidth, or age of computer/software technology. (Curl & Bowers, 2009)

[78] Accessibility is a basic pre-requisite for allowing users to have access to the web page content, while usability provides online users with simple, efficient, rapid and satisfying navigation and interaction. (Buzzi et al., 2009a)

[79] Accessibility is a basic pre-requisite for allowing users to explore web page content, while usability provides online users with simple, efficient, and satisfying navigation and interaction. (Buzzi et al., 2009b)

[80] A website or application is accessible if it can be used by all, including people with disabilities. An accessible (Web) user interface means that potential technical barriers have been eliminated, and thus anyone can interact with it....Accessibility is a basic pre-requisite for allowing users to have access to the web page content, while usability provides online users with simple, rapid and satisfying navigation and interaction. (Akhter et al., 2009)

[81] The term 'Web Accessibility' refers to the creation and development of the web in such a way that web contents are easily accessible by

everyone...Web Accessibility means to develop the web in a way that allows disabled and older people to access and contribute to the web as it would for any normal person. (Talib, Shuqin, Abrar, & Shafiq, 2009)

[82] Web accessibility refers to making the World Wide Web accessible and available to everyone, including people with disabilities and senior citizens. (Alexandru & Alecu, 2010)

[83] Web accessibility is the practice of making Web sites accessible to people who require more than just traditional Web browsers to access the Internet. (Bradbard, Peters, & Caneva, 2010)

[84] Web accessibility refers to persons with disabilities which access the Web content. From this point of view, Web accessibility means a web design that allows people with disabilities to interact with Web pages effectively. (Isaila & Nicolau, 2010)

[85] Web accessibility is an umbrella term for the study of the adequacy of Web technologies to users with special needs such as people with blindness, cognitive disabilities, etc. This adequacy can be viewed from two perspectives: (1) stricter, where accessibility means the ability to access (e.g. a person with blindness cannot grasp information conveyed in images); and (2) broader, where the term represents how easily these users can interact with a Web page. (Lopes & Carrico, 2010)

[86] Web Accessibility means universal access on the Web, regardless the kind of hardware, software, network platform, language, culture,

geographic location and users' capabilities. (Martin, Mazalu, & Cechich, 2010)

[87] Accessibility is making the content of a Website available to everyone, including those with physical disabilities and cognitive learning problems. (Wijayaratne & Singh, 2010)

[88] Web accessibility means ensuring that online content, services or applications can be accessed and used by everyone, including those with special needs. (Leporini, Buzzi, & Buzzi, 2010)

[89] ... means designing a web site that can technically be accessed by users with impairments...An accessible web site means that any user, using any type of assistive technology (such as screen readers, alternative pointing devices or alternative keyboards) can successfully access the content on a web site. (Lazar et al., 2010)

[90] Website accessibility is mainly concerned with easy web content fruition by different categories of people, including those navigating the web through assistive technologies, which provide their users with alternative ways of accessing web pages. (Fogli et al., 2010)

[91] Accessibility is a basic requirement for every system or product in order to guarantee equal access, opportunity and use to all, including the differently-abled. An accessible Web also helps people with disabilities to participate more actively in society. (Buzzi, Buzzi, & Leporini, 2010)

[92] ... a Web application or page is accessible if people with disabilities - including people requiring assistive technologies such as screen readers, screen magnifiers, or speech input - are able to access any information from it and perform any operations it implements. (Trewin et al., 2010)

[93] ...accessibility refers to the fact that something is accessible to users regardless of the means of access and their individual problems or limitations. In the area of information systems consists of a quality attribute that can be described through the development of flexible design solutions to accommodate the diverse needs of a large portion of users, regardless of age, disability or technology. Web-accessibility corresponds to the possibility that any user using any agent (software or hardware that retrieves and serializes Web content), can understand and interact with the content of a website. (Affonso de Lara, Watanabe, dos Santos, & Fortes, 2010)

[94] Accessibility (the property of a website such that “people with some impairment can use it with the same effectiveness as non-disabled people”) deals not only with technicalities of a user interface, but also with the way people perceive, interpret and act on the user interface. (Brajnik, Yesilada, & Harper, 2010)

[95] Web accessibility means overcoming all disabilities that limit its access. It means that people with disabilities can use and perceive, understand, navigate, interact with and contribute to the web. We include all disabilities that affect web access, including visual, auditory, physical,

speech, cognitive and neurological disabilities. (De Andrés, Lorca, & Martínez, 2010)

[96] Web accessibility can be defined as the degree to which a site is accessible to the largest possible range of people. The more people are able to access a website, the more accessible is the site. At its core, Web accessibility emphasizes making website accessible to persons with disabilities and involves removing potential barriers to access caused by inconsiderate website designs. (Latif & Masrek, 2010)

[97] ... accessibility is the ability to access the contents of a web site by all visitors. (Hassanzadeh & Navidi, 2010)

[98] Web accessibility is the practice of making Web sites accessible to all, particularly those with disabilities... Web accessibility is the practice of making Web sites accessible to people who require more than just traditional Web browsers to access the Internet. ... “accessible Web site” is designed to accommodate a wider set of ways students can access a Web site’s content. Many Web sites are designed with visual aesthetics, rather than equal access, as the goal. (Bradbard et al., 2010)

[99] Accessibility is a general term used to describe the degree to which a product is accessible by as many people as possible, and it is often used to focus on people with disabilities and their right of access to entities often through the use of assistive technology. (Wang et al., 2010)

[100] The term “accessibility” generally refers to the application of technical solutions to the design of a web site in order to render it more accessible to users, in particular users of assistive technologies (such as screen reading technology). (Johnson, Rowley, Craven, Johnson, & Butters, 2010)

[101] Web accessibility refers to the ability to use content and services independently of the disability and hardware and software availability. (Belingardi & Obradovic, 2011)

[102] Web accessibility is the concept of making sure that web sites can work properly for users with disabilities that are using alternative input or output devices, such as screen readers or adaptive keyboards. (Lazar et al., 2011)

[103] Web accessibility aims to help people with disabilities to perceive, understand, navigate, interact, and contribute to the web. (Yesilada, Brajnik, & Harper, 2011)

[104] An accessible site is simply a site that disabled people can easily navigate and access all of its contents. Web accessibility should encompass all kind of disability that affect the access to the web including the physical, visual, auditory, speech, and neurological disabilities i.e., low view people should be able to adjust the size of characters, according to their needs for reading. (AlDhaen, El Zant El Kadhi, & Al-Obaidy, 2011)

[105] Accessibility to the websites, refers to the extent by which the website, including the technology such as hypertext coding is barrier free to all users, thus providing enhancements that enable people with disabilities to move towards independence. (Banday & Shah, 2011)

[106] Accessibility of a website refers to the ability of all people to use a website irrespective of their disabilities or the client devices they use to access internet. Accessibility is an important aspect of websites in general and of public websites in particular, to be able to serve all citizens equally. (Abdelgawad, Snaprud, & Krogstie, 2011)

[107] Web accessibility is about making the website accessible for different levels of users and also to people with different levels of abilities and disabilities. Moreover, making the website accessible to all Internet users regardless of the type of the browsing technology they're using is important. (Albalawi, Algosaibi, & Aljohani, 2011)

[108] Web Accessibility is characterized by the possibility of people being able to utilize the Internet and Information Systems, regardless of their physical- motor, perceptual, cultural and social capacities. (Capra, Ferreira, Da Silveira, Ribeiro, & Modesto, 2011)

[109] Accessibility is a prerequisite that permits users to perceive online content and interact, while usability enhances the quality of the interaction, which should be simple, efficient and satisfying. (Mori et al., 2011)

[110] Web accessibility refers to providing equal access and equal opportunity to the Internet for people with disabilities. Web accessibility means that people with disabilities can perceive, understand, navigate, and interact with the Web, and that they can contribute to the Web. (Brobst, 2011)

[111] ... information accessibility corresponds to making it possible for any user, using any user agent (software or hardware designed for viewing Web content) to understand and interact with information and communications technology (ICT) products, regardless of disability, language, or technological constraints. (Yao, Qiu, Huang, Du, & Ma, 2011)

[112] A website is inaccessible when people with perceptual or motor impairments cannot technically use the website. People with perceptual or motor impairments often use assistive technologies, such as screen readers (computer-synthesized speech output, AKA text-to-speech), speech recognition (speech input), and alternative input and output devices. For a website to be considered accessible, it must be flexible enough to work with various input and output devices. It's not that the web pages must have code added for each additional impairment, but instead, when appropriate coding standards and labeling conventions are used, this will make the website accessible for people with various impairments. (Olalere & Lazar, 2011)

[113] Web accessibility aims to help people with disabilities to perceive, understand, navigate, interact with, and contribute to the Web. (Harper & Chen, 2012)

[114] Accessibility in terms of web design generally refers to facilitating the use of technology for people with disabilities with any impairment, not matter what its severity.... The accessibility can be defined as the quality of a web site that makes it possible for people to use it - to find it navigable and understandable - even when they are working under limiting conditions or constraints. (Baowaly & Bhuiyan, 2012)

[115] Website accessibility refers to the practice of making websites accessible to all users inclusive of race, nationality, religion and disability. Website accessibility includes, but is not limited to, the communication style of the text as well as the technical development of the website. (Grantham, Grantham, & Powers, 2012)

[116] Web accessibility refers to the ability to access web and its contents for all people regardless of the disability they have from (physical, cognitive or sensorial disability), or disabilities arising from the use contexts (technological or environmental contexts). (Márquez et al., 2012)

[117] Web accessibility means that people, disregarding of their abilities can access the Web. (Mereuţă, Aupetit, & Slimane, 2012)

[118] With websites, the term traditionally refers to the development of websites accessible to all users who may want to access them, independent of the abilities or disabilities of the users. When websites are correctly designed and developed, all users can have equal access to information and functionality. (Luján-Mora & Masri, 2012)

[119] Web Accessibility is the umbrella term that expresses the process to which a web site is made usable to all visitors, including those with disabilities. More specifically, web accessibility means that people with disabilities can perceive, understand, navigate, and interact with the web, and that they can contribute to the web. In addition to complying with the law, an accessible web site can bring in huge benefits on to web sites and businesses. In other hand developing accessible applications has become a social responsibility for the software engineering industry today. (Wanniarachchi & Jayathilake, 2012)

[120] An accessible Web means that the Web can be used by all, regardless of the impairments users may have. It means it does not have barriers that make the interaction impossible or the content not reachable. A Web page that excludes a user from its service cannot be classified as accessible. (Fernandes, Costa, Duarte, & Carriço, 2012)

[121] Web accessibility means overcoming all disabilities that prejudice access to the web: It means that people with disabilities can use it and perceive, understand, navigate, and interact with the web, and they can contribute to the Web. We include all disabilities that affect web access, including visual, auditory, physical, speech, cognitive, and neurological disabilities. (Lorca, Andrées, & Martínez, 2012)

[122] Web accessibility is about making Web sites accessible and useable by all people. (Kurt, 2012)

[123] Web accessibility is the idea of a “barrier-free” web. To achieve accessibility to the web, web administrators need to enable web content

to be accessible to users at various levels and for different needs. In particular, people with disabilities can access the internet through specially designed equipments, technologies, and resources that are currently available. Web accessibility can be defined as the ability to acquire complete information from the web by anyone – regardless of limited software environment, limited hardware, and disadvantaged capabilities. (Li, Yen, Lu, & Lin, 2012)

[124] Web accessibility relates to the need to enable people with disabilities to use the Web. Web accessibility also relates to older people who may have changing abilities as they age. It is also true that other people besides those with disabilities benefit from a more accessible Web experience, such as people with slow internet connections and temporary disabilities. (Conway, Brown, Hollier, & Nicholl, 2012)

[125] ... web accessibility refers to the access of a website contents by any person regardless of browsing technology. The accessed information shall be fully understandable and user shall be able to interact with website if required. In broader sense, user friendly web designs are said to be accessible designs. The quality of accessible web designs is measured by considering layout, readability, color scheme, browser independency and some special requirements of using adaptive or assistive technologies. (Bakhsh & Mehmood, 2012)

[126] Web accessibility is characterized by people being able to use the Internet regardless of their physical-motor, perceptual, cultural and social capabilities. (Capra, Ferreira, da Silveira, & Ferreira, 2012)

[127] Web accessibility implies that all people – including those with disabilities – can perceive, understand, navigate, and interact with the web, and that they can also contribute to it without barriers. In this context, universal design (a.k.a. design for all, inclusive design) is the process of creating (web)devices, environments, systems, and processes that are usable by people with the wide strange of abilities, operating within the widest possible range of situations. (Kerkmann & Lewandowski, 2012)

[128] An accessible website will be one that has been designed so that people with functional limitations (e.g. visual, motor, cognitive and auditory impairments) and situational limitations (e.g. those using alternative web-access equipment) can freely access the content of the site. An accessible website is "perceivable, operable and understandable", without barriers, for all people. (Parkinson & Olphert, 2010)

[129] Web accessibility encompasses both the technical and service viewpoints of web design. The technical aspect refers to making web sites accessible to people who require more than the usual web browser to access the internet, and the service aspect means ensuring information and services are provided in a way that can be easily accessed and used by people with different physical and mental capabilities. Beyond being able to access information and services lies the ability to use the sites where this information is provided in a way that does not require extraordinary effort. (Ambrozic, Southwell, & Slater, 2012)

[130] ... "Web accessibility" is the possibility that any person accessing the Web in different situations. These situations involve not only technology requirements necessary for the interaction, but also user characteristics such as your skills, preferences, needs and different motor and cognitive limitations. (Dias, de Mattos Fortes, & Masiero, 2012)

[131] Web accessibility is the inclusive practice of making websites usable by people of all abilities and disabilities. (Chevalier et al., 2013)

[132] ....refers to web pages being easily usable by all end users...Web accessibility refers to construction of a web site such that all users can access its information, regardless of their age or physical limitations, and can easily navigate its environment. (Iwata, Kobayashi, Tachibana, Shirogane, & Fukazawa, 2013)

[133] Web accessibility refers the ability to access a website from different browser platforms, either software or hardware related. This should be done in such a way that ensures that government websites are accessible to the target users. (Kituyi & Anjoga, 2013)

[134] Web accessibility refers to the inclusive practice of making web based applications usable by people of all abilities and disabilities. When sites are correctly designed, developed and edited, all users can have equal access to information and functionality. (Anand, Geethamsi, Chary, & Babu, 2013)

[135] By accessibility, it is meant that the web content should be understood, navigated and interacted with to the fullest degree as intended by the

creator of the web content. One group of people that are especially affected by the lack of accessibility are people with disabilities including: visual, auditory, physical, speech, cognitive, or neurological. (Kurt, 2013)

[136] Accessible web sites typically meet the needs of people with perceptual impairments (low vision or blind, deaf or hard of hearing), motor impairments (limited or no use of hands for pointing or typing), and some cognitive impairments. Accessible web sites (which are essentially web sites that are flexible to the user's technology and environment) also tend to increase usability for users of mobile devices. (Lazar et al., 2013)

[137] For a website to be accessible, it must be sufficiently flexible to be used by all of these assistive technologies. [screen readers, alternate keyboards, and refreshable Braille displays] Accessible design is therefore essential to allow people with disabilities to use the Internet more effectively. (Michopoulou & Buhalis, 2013)

[138] Web accessibility means that the web site can be accessed and used effectively by people with and without disabilities. (Doush, 2013)

[139] .... accessibility could be defined as the ease in which people with disabilities, people from different geographic regions and people having different internet connections could access the websites'. (Sambhanthan & Good, 2013)

[140] The ability of a person with disabilities to access a given service or product or execute a given activity in an equal manner as a person who does not have any kind of disability, is the definition of accessibility we adopt in our article. In the world of ICT, the term accessibility can be simply defined as the existence of interfaces that can be used, acknowledged and perceived in the same manner by all users, whether they are disabled or not. (Gonçalves, Martins, Pereira, Oliveira, & Ferreira, 2013)

[141] Web accessibility means that persons with disabilities can use the web on an equal basis with others. (Calle-Jimenez, Sanchez-Gordon, & Luján-Mora, 2014)

[142] ... how easily and effectively a product or service can be accessed and used ... good accessibility is designed for the full range of capabilities, as well as for the context of use or environmental constraints. (Horton & Quesenbery, 2014)

[143] Web Accessibility is the use of Internet resources and access to information without barriers, regardless of cognitive, perceptual or physical capacities of a person. (Modesto & Ferreira, 2014)

[144] Web accessibility means that everyone can benefit from all available information services, regardless of disabilities.... It refers to equitable access to services offered on web sites, regardless of a person's physical health or geographic location. (Park & Lim, 2014)

[145] Web accessibility refers to enabling the differently abled people to use the Web. More specifically, it is about enabling the DAP [Differently Abled People] to perceive, understand, navigate, and interact with the Web. (J. Babu & Sekharaiah, 2014)

[146] Web accessibility means that people can access contents of web pages whichever disabilities they suffer (aging, impairment. . .). (Aupetit & Rouillé, 2014)

[147] By transposing the accessibility concept to the Web environment, it is possible to acknowledge that Web accessibility is the existence of Web interfaces and platforms that can be used and perceived by all users, in an equal manner. (Gonçalves, Martins, & Branco, 2014)

[148] Web accessibility usually refers to creating websites accessible to all users who want to access them, regardless of users' disability. When websites are correctly designed and developed, all users can have access to their information and functionality. ....the objective of the web accessibility is to ensure that people with disabilities can access websites just like everyone else. (Luján-Mora et al., 2014)

[149] Web accessibility means ensuring that anyone including those with disabilities and the elderly can access all information provided by websites in any technical environment without much special skill. (Park et al., 2014)

[150] Web accessibility entails overcoming all disabilities that prejudice Internet access: it means that people with disabilities can use it and

perceive, understand, navigate, and interact with the Web, and they can contribute to the Web. (Martínez, De Andrés, & García, 2014)

[151] Web Accessibility means that people with different types of limitation can perceive, understand, navigate, interact, and contribute with the Web. Accessibility barrier is anything that makes difficult or impossible for people with disability to use the Web. (Santana & Baranauskas, 2014)

[152] eAccessibility is a concept which ensures that all people of all levels of ability have the same access to information made available on the internet as everybody else. This includes people with disabilities and elderly people with reduced functional capabilities. (Huffaker, 2014)

[153] The ability of a given person, with a disability or incapacity, to access a product or service or to execute a task in an equal manner as one without any impairment.... By transposing the accessibility concept to the Web environment, it is possible to acknowledge that Web accessibility is the existence of Web interfaces and platforms that can be used and perceived by all users, in an equal manner. (Gonçalves et al., 2014)

[154] The feature of websites that produces no or minimal obstacles for any users trying to access its contents irrespective of disabilities. A website is inaccessible when people with perceptual or motor impairments cannot technically use the website. (Karkin & Janssen, 2014)

[155] Web accessibility encompasses all disabilities that affect access to the Web, including visual, auditory, physical, speech, cognitive, and neurological disabilities. While access to people with disabilities is the primary focus of web accessibility, it also benefits people without disabilities. Thus, accessible technology is technology that users can adapt to meet their visual, hearing, dexterity, cognitive, and speech needs and interaction preferences. (Sánchez-Gordón & Moreno, 2014)

[156] …accessibility refers to the viability of an individual with disabilities to access and use information as it is presented on the public library's website. Accessibility considers whether information can be read by manipulating text on the screen (enlarging text size, changing color and contrast) or through the use of other adaptive technologies, such as screen readers or refreshable braille displays. (Maatta Smith, 2014)

[157] Accessibility corresponds to the right of any person be able to enjoy products, services and information that belongs to life in society, regardless of their physical and motor, perceptual, cultural and social skills. One of the steps to promote that accessibility is through the removal of barriers that hinder the daily activities. In the context of web pages, e-accessibility is the capacity of interaction and understanding of anyone using any kind of navigation technology for access to information. (Pereira, Ferreira, Braga, de Castro Salgado, & Nunes, 2014)

Appendix C. Empirical evaluation of the concurrent and retrospective verbal protocol for blind and sighted users (Study 2) – material

This section presents the material used in study 2 (Chapter 4) of this thesis.

Introductory script

Hello, my name is Andreas Savva and I will be running this session with you today.

Today we are going to spend about 2 hours doing some tasks on different websites. There are 4 different websites. On each of these websites you will be asked to do a couple of tasks that are fairly typical of the types of things people try to do. Please do the tasks as you would at home or your office.

On two of the websites, I'm going to ask you to think-aloud about what you are trying to do on the website while you do it. Each time you encounter a problem of any kind, I would like you to rate it on a scale of 1 (cosmetic problem only), 2 (minor problem), 3 (major problem) and 4 (catastrophic problem). We asked you to rate the problems on scale from 1 to 4, which is the usual rating scale evaluators' use when evaluating websites with users.

On the other two websites, I'm going to have you just undertake tasks as you normally would. No thinking aloud will be needed. At the end of each of those tasks, we will play the video of you doing it. While the video is playing, I will ask you to think aloud about what was happening during the task. Each time you encounter a problem, I would like you to rate it in the same way.

After each set of websites, I will ask you to complete a couple of questions about the method we used.

If you fall quiet for a little bit during any of these think aloud sessions, I'll prompt you with something like "What are you thinking about?" just to remind you to continue to vocalize what is happening in the task.

Before we start either set of tasks, I'll give you a longer demonstration of what you are being asked to do.

One thing I want to emphasise is that we are not testing you or your abilities. You are helping us test these websites. If you ever feel that you are lost or cannot complete a task with the information that you have been given, please let me know. You can also stop at any time.

Informed consent form

PhD study on two different think aloud methods

This study is part of my PhD research. It is an investigation into two different think aloud methods that are very often used for evaluating web sites. The think aloud during the task and think aloud during the replay of the task.

Before you participate in this study please complete Section A, printing your name in the first space and then sign at the end.

Once the study is over and you have been debriefed, you will be asked to initial the two statements in Section B, to indicate your agreement.

Section A

I, ________________________________, voluntarily give my consent to participate in this study for the investigation into the two different think aloud methods and be recorded. I have been informed about, and feel that I understand the basic nature of the project. I understand that I may withdraw from the study at any time without prejudice. I also understand that my information is confidential and recording is for research purposes only. Only Andreas Savva, Professor Helen Petrie and Dr. Chris Power will have access to the data collected today in its original format and it will only be shared with other researchers working on web accessibility in the Human Computer Interaction Research Group at the University of York in an anonymous format.

______________________________ ___________________
Signature of Research Participant Date

Section B

Please initial each of the following statements when the study has been completed and you have been debriefed.

I have been adequately debriefed	Your initials:
All my questions have been answered	Your initials:

Demographics questions

This questionnaire was provided to blind users. For sighted users, the same questionnaire was used without questions 8 to 14.

Demographic Information:

1. Sex
    a. Male
    b. Female

2. What is your highest level of education?
    a. Secondary School
    b. Undergraduate degree
    c. Masters degree
    d. Doctoral degree
    e. Other:

    (if b,c,d) What is your highest-degree in?

3. Are you currently a student?
    a. Yes

        What level of qualification are you taking?

        ____

        What are you currently studying?

        ____

    b. No

4. What is your age? _________________

5. How would you rate your level of experience with the Web?

+ ------- + ------- +------- + ------- +

Very low Very high

6. How would you rate your level of expertise with the Web?

+ ------- + ------- +------- + ------- +

Very low Very high

7. What browser(s) do you typically use to browse the Web? (Collect name and version)
____
____

8. What is your sight status?
____
____

9. Has your sight status always been the same?
____
____

10. Do you use assistive technology/technologies to browse the Web?
    a. Yes
    b. No (if no end survey)

11. What assistive technology/technologies do you use to browse the Web?
    a. Home:
        i. Software: ____________________________
           version: ____________________________
        ii. Software: ____________________________
           version: ____________________________

iii. Software: ______________________________
version: ______________________________

b. Work:
i. Software: ______________________________
version: ______________________________
ii. Software: ______________________________
version: ______________________________
iii. Software: ______________________________
version: ______________________________

12. How long have you used your assistive technology/technologies?
____
____

13. What is your experience with your assistive technology/technologies?

Very limited	Limited	Adequate	Good	Very Good

14. What is your expertise with your assistive technology/technologies?

Very limited	Limited	Adequate	Good	Very Good

Debriefing

Thank you very much for participating in the study.

This study was an investigation into two different think aloud methods that are very often used for evaluating web sites. These were the think aloud during the task and think aloud during the replay of the task. We want to find which method can identify more problems overall on a website, which is easier and more pleasant for the participants to do, which finds the more problems in the least time.

We also wanted to find out how much extra effort doing the two method puts on people and whether that interrupts and distracts them. The main way we measured the extra effort was by asking you to do the complicated questionnaire, which was the NASA Task Load Index, known as the NASA TLX. The NASA TLX has the measure of the overall effort, or workload of the task, but also six different measures of different kinds of effort, the mental demand, physical demand, temporal demand, effort, frustration and performance of the user.

Also, we wanted to check if thinking aloud and doing the task interrupts the flow of the task, your concentration or made you self-conscious about what you were doing, as well as how it differs from doing the task in real life.

In the study we are asking both sighted people and blind people to try out the two methods, as we are interested in which of the two methods is better for both sighted and blind people.

The results of this study will help us decide which is the best method for evaluating web sites.

Thank you again for you participation. Do you have any further questions about the study?

Thank you again, we need to complete Section B of the Consent form now.

Appendix D. Empirical study of the problems between blind and sighted users on the web (Study 3) – material

This section presents the material used in study 3 (Chapter 5) of this thesis. The demographics questionnaire used was the same as in study 2 (Appendix C.).

Introductory script

PhD Study on problems users encounter with on the web

This study is part of PhD research. It is an investigation into the types of problems users encounter with on the web in order identify and characterise what are the types of problems. This will provide us with an insight regarding the problems.

With your permission we will record the session, so that we can study the problems you encounter in details afterwards. Only me, Helen Petrie and Chris Power, as researchers working on web accessibility in the HCI group at the University of York will be allowed to view the recording.

You will be asked to do a number of tasks on three websites, which are fairly similar to the types of things people usually do. Please do the tasks as you would at home or your office. Please do the tasks in silence. At the end of each task we will play a video of you doing the task and you can do the think-aloud on what was happening during the task. If you fall quiet for a bit during the think aloud session, when we play the video, I will prompt you with something like “What are you thinking about?” just to remind you to continue to vocalise your thoughts.

Each time you encounter a problem of any kind, I would like you to detailed describe the problem.

After completing the task on each website, I will ask you to complete a questionnaire about the website. At the end of all the three websites, I will ask you to complete a questionnaire about the whole session and some demographic questions.

One thing I want to emphasise, is that we are not testing you or your abilities. You are helping us to test these websites. If you ever feel that you are lost or cannot complete a task with the information that you have been given, please let me know. You can also stop at any time.

Before you participate in this study, please complete Section A of the consent form, printing your name in the first space and then signing at the end.

Once the study is over and you have been debriefed, you will be asked to initial the two statements in Section B, to indicate your agreement.

Informed consent form

Section A

I, ________________________________, voluntarily give my consent to participate in this study on the problems users encounter with on the web and be recorded. I have been informed about, and feel that I understand the basic nature of the project. I understand that I may withdraw from the study at any time without prejudice. I also understand that my information is confidential and recording is for research purposes only. Only Andreas Savva, Professor Helen Petrie and Dr. Chris Power will have access to the data collected today in its original format and it will only be shared with other researchers working on web accessibility in the Human Computer Interaction Research Group at the University of York in an anonymous format.

______________________________ ___________________

Signature of Research Participant Date

Section B

Please initial each of the following statements when the study has been completed and you have been debriefed.

I have been adequately debriefed Your initials:

All my questions have been answered Your initials:

Debriefing

Thank you very much for participating in the study.

This study was an investigation into the problems users encounter with, on the web. We want to find more information about the problems, the type of problem, and why these problems occur. This will give us insights in order to help us improve the web.

In this study, we are asking both blind and sighted participants to perform the tasks, as we are interested in comparing the problems that come across, and whether they differ.

Thank you again for your participation. Do you have any further questions about the study?

Thank you again, we need to complete Section B of the Consent form now.

Appendix E. Empirical study of the benefits of specific design solutions on blind users' experience in search and browse websites (Study 5) – material

This section presents the material used for study 5 (Chapter 8) of this thesis. The demographics questionnaire is the same one used in study 2 (Appendix C.).

Introductory script

This study is part of a PhD research. We are going to spend about 2 hours doing some tasks on different websites. There are 3 different websites. On each of these websites you will be asked to do a couple of tasks that are fairly typical of the types of things people try to do. Please do the tasks as you would at home or your office.

After completing all the tasks on each website, I will ask you to complete a questionnaire about the whole session. At the end of all three websites I will ask you to complete some demographic questions.

With your permission, we will record the session. Only me, Helen Petrie and Chris Power, as researchers working on web accessibility in the HCI group at the University of York will be allowed to view the recordings.

One thing I want to emphasize, is that we are not testing you or your abilities. You are helping us to test these websites. If you ever feel that you are lost or cannot complete a task with the information that you have been given, please let me know. You can stop at any time.

Before you participate in this study, please complete Section A, printing your name in the first space and then signing at the end. Once the study is over and you have been debriefed, you will be asked to initial two statements in Section B, to indicate your agreement.

Informed consent Form

Section A

I, ________________________________, voluntarily give my consent to participate in this study on evaluation of different websites and be recorded. I have been informed about, and feel that I understand the basic nature of the project. I understand that I may withdraw from the study at any time without prejudice. I also understand that my information is confidential and recording is for research purposes only. Only Andreas Savva, Professor Helen Petrie and Dr. Christopher Power will have access to the data collected today in its original format and it will only be shared with other researchers working on web accessibility in the Human Computer Interaction Research Group at the University of York in an anonymous format. I will compensate £15 per hour of evaluation in gift voucher for participating in the study.

______________________________ ___________________

Signature of Research Participant Date

Section B

Please initial each of the following statements when the study has been completed and you have been debriefed.

I have been adequately debriefed Your initials:

All my questions have been answered Your initials:

NASA TLX Questionnaire

Rating Scales Instructions

We are not only interested in assessing your performance but also the experiences you had during the different task conditions. Right now we are going to describe the technique that will be used to examine your experiences. In the most general sense we are examining the “workload” you experienced. Workload is a difficult concept to define precisely, but a simple to understand generally. The factors that influence your experience of workload may come from the task itself, your feelings about your own performance, how much effort you put in, or the stress and frustration you felt. The workload contributed by different task elements may change as you get more familiar with a task, perform easier or harder version of it, or move from one task to another. Physical components of workload are relatively easy to conceptualize and evaluate. However, the mental components of workload may be more difficult to measure.

Since workload is something that is experienced individually by each person, there are no effective “rules” that can be used to estimate the workload of different activities. One way to find out about workload is to ask people to describe the feelings they experienced. Because workload may be caused by many different factors, we would like you to evaluate several of them individually rather than lumping them into a single global evaluation of overall workload. This set of six rating scales was developed to use in evaluating experienced during different tasks. Please listen the descriptions of the scales carefully. If you have a question about any of the scales, please ask me about it. It is extremely important that may be clear to you. You may ask me again about the description of each scale.

After performing each of the tasks, you will be given a sheet of rating scales.

You will evaluate the tasks using the scale 1 to 20, pointing the scale that matches your experience. Each scale goes from “1 = low” to “20 = high”. Note that “own performance” goes from “1 = good” to “20 = poor”.

This order has been confusing for some people. Please consider your responses carefully in distinguishing among the different task conditions. Consider each scale individually. Your ratings will play an important role in the evaluation being conducted. Thus, your active participation is essential to the success of this experiment and is greatly appreciated by all of us.

Sources of Workload Evaluation Instructions

Throughout this experiment the rating scales are used to assess your experiences in the different task conditions. Scales of this sort are extremely useful, but their utility suffers from the tendency people have to interpret in individual ways. For example, some people feel that mental or temporal demand are the essential aspects of workload regardless of the effort they expended on a given task or the level of performance they achieved. Others feel that if they performed well the workload must have been low and if they performed badly it must have been high. Yet others feel that effort or feelings of frustration are the most important factors in workload and so on. The results of previous studies have already found every conceivable pattern of values. In addition, the factors that create levels of workload differ depending on the task. For example, some tasks might be difficult because they must be completed very quickly. Others may seem easy or hard because of the intensity of mental or physical effort required. Yet others feel difficult because they cannot be performed well, no matter how much effort is expended.

The evaluation you are about to perform is a technique that has been developed by NASA to assess the relatively importance of six factors in determining how much workload you experienced. The procedure is simple: You will be presented with a series of pairs of rating scale titles (for example, Effort vs Mental Demands) and asked to choose which of the items was more important to your experience of workload in the task(s) that you just performed.

Please tell me the Scale Title that represents the more important contributor to workload for the tasks you performed in this experiment.

After you have finished the entire series you will be able to use the pattern of your choices to create a weighted combination of the ratings from that task into a summary workload score. Please consider your choices carefully and make them consistent with how you used the rating scales during the particular task you were asked to evaluate. Don't think that there is any correct pattern: we are only interested in your opinions.

If you have any questions, please ask them now. Otherwise, start whenever you are ready. Thank you for your participation.

Pairwise Comparisons

Effort	or	Performance
Temporal Demand	or	Frustration
Temporal Demand	or	Effort
Physical Demand	or	Frustration
Performance	or	Frustration
Physical Demand	or	Temporal Demand
Physical Demand	or	Performance
Temporal Demand	or	Mental Demand
Frustration	or	Effort
Performance	or	Mental Demand
Performance	or	Temporal Demand
Mental Demand	or	Effort
Mental Demand	or	Physical Demand
Effort	or	Physical Demand

Frustration or Mental Demand

Rating Sheet

Mental demand (1 = low to 20 = high)

How much mental and perceptual activity was required (e.g., thinking, deciding, calculating, remembering, looking, searching, etc.)? Was the task easy or demanding, simple or complex, exacting or forgiving?

Low High

Physical demand (1 = low to 20 = high)

How much physical activity was required (e.g., pushing, pulling, turning, controlling, activating, etc.)? Was the task easy or demanding, slow or brisk, slack or strenuous, restful or laborious?

Low High

Temporal demand (1 = low to 20 = high)

How much time pressure did you feel due to the rate or pace at which the tasks or task elements occurred? Was the pace slow and leisurely or rapid and frantic?

Low High

Performance (1 = good to 20 = poor)

How successful do you think you were in accomplishing the goals of the task set by the experimenter (or yourself)? How satisfied were you with your performance in accomplishing these goals?

Low High

Effort (1 = low to 20 = high)

How hard did you have to work (mentally and physically) to accomplish your level of performance?

Low High

Frustration level (1 = low to 20 = high)

How insecure, discouraged, irritated, stressed and annoyed versus secure, gratified, content, relaxed and complacent did you feel during the task?

Low High

Debriefing

Thank you very much for participating in the study.

This study was an investigation into the benefits of different website designs on the filtering options of websites. This will provide us with knowledge whether these designs can help us to improve users' experience on the search systems on the web.

We also wanted to find out how much extra effort doing the tasks in each design puts on people. The main way we measured the extra effort was by asking you to do the complicated questionnaire, which was the NASA Task Load Index, known as the NASA TLX. The NASA TLX has the measure of the overall effort, or workload of the task, but also six different measures of different kinds of effort, the mental demand, physical demand, temporal demand, effort, frustration and performance of the user.

Also, we wanted to check if either of the designs makes users' confident that they completed the task successfully as well as how clear was it what was happening in the page was.

The results of this study will provide us with knowledge whether these designs can help us to improve the web.

Thank you again for your participation. Do you have any further questions about the study?

Thank you again, we need to complete Section B of the Consent form now.

Appendix F. Empirical confirmation study of the benefits of specific design solutions on blind users' experience in an exploratory search (Study 6) - material

This section presents the material used for study 6 (Chapter 9) of this thesis. The demographics questionnaire is the same one used in study 2 (Appendix C.). The informed consent form, NASA TLX questionnaire and debriefing material used in this study are the same as in study 5 (Appendix E.).

Introductory script

This study is part of a PhD research. We are going to spend about two hours doing some tasks on different websites. There are two different websites. On each website, you will be asked to do a couple of tasks that are fairly typical of the types of things people try to do. Please do the tasks as you would at home or your office.

After completing all the tasks on each website, I will ask you to complete some questionnaires about the whole session. At the end of both websites, I will ask you to complete some demographic questions.

With your permission, we will record the session. Only me, Helen Petrie and Chris Power, as researchers working on web accessibility in the HCI group at the University of York will be allowed to view the recordings.

One thing I want to emphasize, is that we are not testing you or your abilities. You are helping us to test these websites. If you ever feel that you are lost or cannot complete a task with the information that you have been given, please let me know. You can stop at any time.

Before you participate in this study, please complete Section A, printing your name in the first space and then signing at the end. Once the study is over and you have been debriefed, you will be asked to initial two statements in Section B, to indicate your agreement.

References

Abanumy, A., Al-Badi, A., & Mayhew, P. (2005). e-Government Website accessibility: in-depth evaluation of Saudi Arabia and Oman. *The Electronic Journal of e-Government, 3*(3), 99-106.

Abascal, J., Arrue, M., Garay, N., & Tomás, J. (2003). A Web Service for Automatic Accessibility Analysis of Web Pages Based on the Use of XML Structures. In: *Proceedings of 10th International Conference on Human-Computer Interaction, Lawrence Erlbaum Associates*, 925-929.

Abdelgawad, A. A., Snaprud, M. H., & Krogstie, J. (2011). Accessibility of norwegian municipalities websites: A decision support tool. In: *Fifth UKSim European Symposium on Computer Modeling and Simulation (EMS)*, 225-230.

Abu-Doush, I., Bany-Mohammed, A., Ali, E., & Al-Betar, M. A. (2013). Towards a more accessible e-government in Jordan: an evaluation study of visually impaired users and Web developers. *Behaviour & Information Technology, 32*(3), 273-293.

Affonso de Lara, S. M., Watanabe, W. M., dos Santos, E. P. B., & Fortes, R. P. (2010). Improving WCAG for elderly web accessibility. In: *Proceedings of the 28th ACM international conference on design of communication*, 175-182.

Aizpurua, A., Harper, S., & Vigo, M. (2016). Exploring the relationship between web accessibility and user experience. *International Journal of Human-Computer Studies, 91*, 13-23.

Akhter, F., Buzzi, M. C., Buzzi, M., & Leporini, B. (2009). Conceptual framework: How to engineer online trust for disabled users. In: *Proceedings of the 2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology-Volume 03*, 614-617.

Albalawi, T., Algosaibi, A., & Aljohani, K. (2011). WAESM: Web Accessibility Educational Specific Model. In: *Proceedings of the 11th International Conference Software Engineering Research and Practice*, 529-533.

AlDhaen, F., El Zant El Kadhi, N., & Al-Obaidy, H. (2011). Accessible web-based educational system. In: *Proceedings of the 12th WSEAS international conference on Neural networks, fuzzy systems, evolutionary computing & automation*, 42-48.

Alexander, D. (2004). WebWatch: How Accessible Are Australian University Web Sites? *Ariadne*(38).

Alexandru, A., & Alecu, C. S. (2010). Introducing engineering students to telemedicine. Case study. In: *Proceedings of the 7th WSEAS international conference on Engineering education*, 465-469.

Alshammari, T., Alhadreti, O., & Mayhew, P. (2015). When to ask participants to think aloud: A comparative study of concurrent and retrospective think-aloud methods. *International Journal of Human Computer Interaction, 6*(3), 48-64.

Ambrozic, K., Southwell, K. L., & Slater, J. (2012). Accessibility of digital special collections using screen readers. *Library hi tech, 30*(3), 457-471.

Anand, G., Geethamsi, S., Chary, R., & Babu, C. M. (2013). Email Access By Visually Impaired. In: *International Conference on Communication Systems and Network Technologies (CSNT)* 597-601.

Andrés, J. D., Lorca, P., & Martínez, A. B. (2009). Economic and financial factors for the adoption and visibility effects of Web accessibility: The case of European banks. *Journal of the American society for Information science and technology, 60*(9), 1769-1780.

Aupetit, S., & Rouillé, V. (2014). Annotation tool for the smart web accessibility platform. In: *International Conference on Computers for Handicapped Persons*, 93-100.

Baazeem, I. S., & Al-Khalifa, H. S. (2015). Advancements in web accessibility evaluation methods: how far are we? In: *Proceedings of the 17th International Conference on Information Integration and Web-based Applications & Services*, 1-5.

Babu, J., & Sekharaiah, C. (2014). A Panorama of Web Accessibility. *International Journal of Computer Science and Mobile Computing, 3*, 311-317.

Babu, R. (2013). Understanding challenges in non-visual interaction with travel sites: An exploratory field study with blind users. *First Monday, 18*(12).

Babu, R., & Singh, R. (2009). Evaluation of web accessibility and usability from blind user's perspective: the context of online assessment. In: *Proceedings ofThe Fifteenth Americas Conference on Information Systems*.

Babu, R., & Singh, R. (2013a). Accessibility and usability of social media: convergence between blind users and design standards. In: *Proceedings of the Nineteenth Americas Conference on Information Systems*, 15-17.

Babu, R., & Singh, R. (2013b). Enhancing learning management systems utility for blind students: A task-oriented, user-centered, multi-method evaluation technique. *Journal of Information Technology and Education: Research, 12*.

Baguma, R., & Lubega, J. T. (2008). A web design framework for improved accessibility for people with disabilities (WDFAD). In: *Proceedings of the 2008 international cross-disciplinary conference on Web accessibility (W4A)*, 134-140.

Baguma, R., Stone, R. G., Lubega, J. T., & van der Weide, T. P. (2009). Integrating accessibility and functional requirements. In: *International Conference on Universal Access in Human-Computer Interaction*, 635-644.

Bailey, J., & Burd, E. (2005). Web accessibility evolution in the United Kingdom. In: *Seventh IEEE International Symposium on Web Site Evolution, 2005.(WSE 2005).* , 79-84.

Bakhsh, M., & Mehmood, A. (2012). Web accessibility for disabled: a case study of government websites in Pakistan. In: *10th International Conference on Frontiers of Information Technology (FIT)*, 342-347.

Banday, M. T., & Shah, N. A. (2011). Challenges of CAPTCHA in the accessibility of Indian regional websites. In: *Proceedings of the Fourth Annual ACM Bangalore Conference*, 31.

Baowaly, M. K., & Bhuiyan, M. (2012). Accessibility analysis and evaluation of Bangladesh government websites. In: *International Conference on Informatics, Electronics & Vision (ICIEV)*, 46-51.

Batra, S. (2009). E-campaigning: The present and future. In: *International Conference on e-Democracy*, 89-98.

Belingardi, G., & Obradovic, J. (2011). Implementation of advanced communication systems and distance learning concept at Politecnico di Torino within the university education perspective. In: *Proceedings of the XV International Scientific Conference on Industrial Systems (IS'11)*, 173.

Berners-Lee, T. (1997). W3C leads program to make the web accessible for people with disabilities (press release). Retrieved from http://www.w3.org/Press/WAI-Launch.html

Bevan, N., Carter, J., & Harker, S. (2015). ISO 9241-11 revised: What have we learnt about usability since 1998? In: *International Conference on Human-Computer Interaction*, 143-151.

Bigham, J. P., Cavender, A. C., Brudvik, J. T., Wobbrock, J. O., & Ladner, R. E. (2007). WebinSitu: a comparative analysis of blind and sighted browsing behavior. In: *Proceedings of the 9th international ACM SIGACCESS conference on Computers and accessibility*, 51-58.

Bolchini, D., Colazzo, S., & Paolini, P. (2006). Requirements for aural web sites. In: *Eighth IEEE International Symposium on Web Site Evolution, 2006. WSE'06.* , 75-82.

Boren, T., & Ramey, J. (2000). Thinking aloud: reconciling theory and practice. *IEEE transactions on professional communication, 43*(3), 261-278.

Borrino, R., Furini, M., & Roccetti, M. (2009). Augmenting social media accessibility. In: *Proceedings of the 2009 International Cross-Disciplinary Conference on Web Accessibililty (W4A)*, 54-57.

Bowers, V. A., & Snyder, H. L. (1990). Concurrent versus retrospective verbal protocol for comparing window usability. In: *Proceedings of the Human Factors Society Annual Meeting*, 1270-1274.

Bradbard, D. A., Peters, C., & Caneva, Y. (2010). Web accessibility policies at land-grant universities. *The Internet and Higher Education, 13*(4), 258-266.

Brajnik, G. (2006). Web accessibility testing: when the method is the culprit. In: *International Conference on Computers for Handicapped Persons*, 156-163.

Brajnik, G. (2008). Beyond conformance: the role of accessibility evaluation methods. In: *International Conference on Web Information Systems Engineering*, 63-80.

Brajnik, G., Yesilada, Y., & Harper, S. (2010). Testability and validity of WCAG 2.0: the expertise effect. In: *Proceedings of the 12th international ACM SIGACCESS conference on Computers and accessibility*, 43-50.

Brebner, M., & Parkinson, M. (2006). The accessibility of New Zealand public library Web sites to people who are blind or vision impaired. In: *LIANZA Conference, Wellington, New Zealand*.

Brobst, J. L. (2011). Access denied: Improving federal policy on Web accessibility. *Proceedings of the American Society for Information Science and Technology, 48*(1), 1-4.

Brophy, P., & Craven, J. (2007). Web accessibility. *Library trends, 55*(4), 950-972.

Bruun, A., & Stage, J. (2015). An empirical study of the effects of three think-aloud protocols on identification of usability problems. In: *Human-Computer Interaction – INTERACT 2015 Lecture Notes in Computer Science*, 159-176.

BSI. (2010). Web accessibility - Code of practice (BS 8878:2010).

Buzzi, M. C., Buzzi, M., & Leporini, B. (2010). Accessibility and usability of web content and applications. *Quality and Communicability for Interactive Hypermedia Systems: Concepts and Practices for Design: Concepts and Practices for Design*, 64-90.

Buzzi, M. C., Buzzi, M., Leporini, B., & Akhter, F. (2009a). Usability and accessibility of ebay by screen reader. In: *Symposium of the Austrian HCI and Usability Engineering Group*, 500-510.

Buzzi, M. C., Buzzi, M., Leporini, B., & Akhter, F. (2009b). User trust in ecommerce services: perception via screen reader. In: *New Trends in Information and Service Science, 2009. NISS'09. International Conference on*, 1166-1171.

Buzzi, M. C., Buzzi, M., Leporini, B., & Senette, C. (2010). Electronic Commerce “in the dark”. In: *International Workshop on Human-Computer Interaction, Tourism and Cultural Heritage*, 12-22.

Byerley, S. L., & Beth Chambers, M. (2002). Accessibility and usability of Web-based library databases for non-visual users. *Library hi tech, 20*(2), 169-178.

Cairns, P. (2007). HCI... not as it should be: inferential statistics in HCI research. In: *Proceedings of the 21st British HCI Group Annual Conference on People and Computers: HCI... but not as we know it-Volume 1*, 195-201.

Caldwell, B., Cooper, M., Guarino Reid, L., & Vanderheiden, G. (2008). Web Content Accessibility Guidelines (WCAG) 2.0. Retrieved from https://www.w3.org/TR/WCAG20/

Calle-Jimenez, T., Sanchez-Gordon, S., & Luján-Mora, S. (2014). Web accessibility evaluation of massive open online courses on geographical information systems. In: *Global Engineering Education Conference (EDUCON), 2014 IEEE*, 680-686.

Capra, E. P., Ferreira, S. B. L., da Silveira, D. S., & Ferreira, A. O. (2012). Evaluation of Web accessibility: an approach related to functional illiteracy. *Procedia Computer Science, 14*, 36-46.

Capra, E. P., Ferreira, S. B. L., Da Silveira, D. S., Ribeiro, B. B., & Modesto, D. M. (2011). Evaluation of web accessibility from the perspective of functional illiteracy. In: *Proceedings of the 10th Brazilian Symposium on Human Factors in Computing Systems and the 5th Latin American Conference on Human-Computer Interaction*, 280-288.

Chandrashekar, S., Stockman, T., Fels, D., & Benedyk, R. (2006). Using think aloud protocol with blind users:: a case for inclusive usability evaluation methods. In: *Proceedings of the 8th* *international ACM SIGACCESS conference on Computers and accessibility*, 251-252.

Chen, Y.-L., Chen, Y.-Y., & Shao, M. (2006). 2005 accessibility diagnosis on the government web sites in Taiwan, R.O.C. In: *Proceedings of the 2006 international cross-disciplinary workshop on Web accessibility (W4A): Building the mobile web: rediscovering accessibility?*, 132-142.

Chevalier, A., Dommes, A., & Martins, D. (2013). The effects of ageing and website ergonomic quality on internet information searching. *Ageing and Society, 33*(06), 1009-1035.

Chisholm, W., & May, M. (2008). *Universal design for web applications: Web applications that reach everyone*: " O'Reilly Media, Inc.".

Chisholm, W., Vanderheiden, G., & Jacobs, I. (1999). Web content accessibility guidelines 1.0. Retrieved from https://www.w3.org/TR/WCAG10/

Conway, V., Brown, J., Hollier, S., & Nicholl, C. (2012). Website accessibility: a comparative analysis of Australian national and state/territory library websites. *The Australian Library Journal, 61*(3), 170-188.

Cooper, M. (2016). Wcag 2.1 under exploration, comments requested by 1 november. Retrieved from https://www.w3.org/blog/2016/10/wcag-2-1-under-exploration/

Coursaris, C. K., Swierenga, S. J., & Whitten, P. (2014). Usability and Accessibility of E-Health Websites: Enabling Nutrition and Physical. *Evaluating Websites and Web Services: Interdisciplinary Perspectives on User Satisfaction*, 204-223.

Coyne, K. P., & Nielsen, J. (2001). Beyond ALT text: making the web easy to use for users with disabilities. *Fremont, California: Nielsen-Norman Group*.

Craven, J. (2003). Access to electronic resources by visually impaired people. *Information Research, 8*(4).

Craven, J. (2006). Web accessibility: A review of research and initiatives. In: *World Library and Information Congress: 72nd IFLA General Conference and Council*.

Craven, J., & Nietzio, A. (2007). A task-based approach to assessing the accessibility of web sites. *Performance measurement and metrics, 8*(2), 98-109.

Curl, A. L., & Bowers, D. D. (2009). A longitudinal study of website accessibility: Have social work education websites become more accessible? *Journal of Technology in Human Services, 27*(2), 93-105.

Curran, K., Walters, N., & Robinson, D. (2007). Investigating the problems faced by older adults and people with disabilities in online environments. *Behaviour & Information Technology, 26*(6), 447-453.

De Andrés, J., Lorca, P., & Martínez, A. B. (2010). Factors influencing web accessibility of big listed firms: an international study. *Online Information Review, 34*(1), 75-97.

De Lima, S. T., Lima, F., & De Oliveira, K. M. (2007). Towards Metrics for Web Accessibility Evaluation. In: *Proceedings of Workshop on Design & Evaluation of e-Government Applications and Services (DEGAS'2007)*, 11-14.

De Lima, S. T., Lima, F., & De Oliveira, K. M. (2009). Evaluating the Accessibility of Websites to Define Indicators in Service Level Agreements. In: *International Conference on Enterprise Information Systems*, 858-869.

De Oliveira Junior, E. A., Motti, V. G., Freire, A. P., & De Mattos Fortes, R. P. (2007). Supporting web page accessibility by using earl reports. In: *Proceedings of the XIII BrazilianSymposium on Multimedia and the Web (Webmedia)*, 32-36.

Department for Work and Pensions Office for Disability Issues and The Rt Hon Mark Harper. (2014). High street could be boosted by £212 billion 'purple pound' by attracting disabled people and their families. Retrieved from https://www.gov.uk/government/news/high-street-could-be-boosted-by-212-billion-purple-pound-by-attracting-disabled-people-and-their-families

Di Lucca, G. A., Fasolino, A. R., & Tramontana, P. (2005). Web site accessibility: Identifying and fixing accessibility problems in client page code. In: *Seventh IEEE International Symposium on Web Site Evolution (WSE 2005).* , 71-78.

Dias, A. L., de Mattos Fortes, R. P., & Masiero, P. C. (2012). Increasing the quality of web systems: by inserting requirements of accessibility and usability. In: *Quality of Information and Communications Technology (QUATIC), 2012 Eighth International Conference on the*, 224-229.

Diggs, J., McCarron, S., Cooper, M., Schwerdtfeger, R., & Craig, J. (2017). Accessible Rich Internet Applications (WAI-ARIA) 1.1. Retrieved from https://www.w3.org/TR/wai-aria-1.1/

Disability Rights Commission. (2004). *The Web: Access and Inclusion for Disabled People; a Formal Investigation*: The Stationery Office.

Doush, I. A. (2013). Web Accessibility for Persons with Motor Limitations. *Disability Informatics and Web Accessibility for Motor Limitations*, 234-262.

Dumas, J. S., & Redish, J. (1999). *A practical guide to usability testing*: Intellect books.

Elo, S., & Kyngäs, H. (2008). The qualitative content analysis process. *Journal of advanced nursing, 62*(1), 107-115.

Equality Act. (2010). London: HMSO.

Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological review, 87*(3), 215-251.

Ericsson, K. A., & Simon, H. A. (1993). Protocol analysis. Verbal protocols as data: Cambridge, MA: MIT Press.

Federici, S., Micangeli, A., Ruspantini, I., Borgianni, S., Corradi, F., Pasqualotto, E., & Olivetti Belardinelli, M. (2005). Checking an integrated model of web accessibility and usability evaluation for disabled people. *Disability and Rehabilitation, 27*(13), 781-790.

Fernandes, N., Costa, D., Duarte, C., & Carriço, L. (2012). Evaluating the accessibility of web applications. *Procedia Computer Science, 14*, 28-35.

Fogli, D., Colosio, S., & Sacco, M. (2010). Managing accessibility in local e-government websites through end-user development: a case study. *Universal Access in the Information Society, 9*(1), 35-50.

Freire, A. P. (2012). *Disabled people and the Web: User-based measurement of accessibility.* PhD thesis, University of York.

Freire, A. P., Bittar, T. J., & Fortes, R. P. (2008). An approach based on metrics for monitoring web accessibility in Brazilian municipalities web sites. In: *Proceedings of the 2008 ACM symposium on Applied computing*, 2421-2425.

Freire, A. P., Fortes, R. P., Turine, M. A., & Paiva, D. (2008). An evaluation of web accessibility metrics based on their attributes. In: *Proceedings of the 26th annual ACM international conference on Design of communication*, 73-80.

Freire, A. P., Russo, C. M., & Fortes, R. (2008). The perception of accessibility in web development by academy, industry and government: a survey of the brazilian scenario. *New Review of Hypermedia and Multimedia, 14*(2), 149-175.

Freire, A. P., Russo, C. M., & Fortes, R. P. (2008). A survey on the accessibility awareness of people involved in web development projects in Brazil. In: *Proceedings of the 2008 international cross-disciplinary conference on Web accessibility (W4A)*, 87-96.

Friedel, M. J., & Wood, D. (2008). Challenging design students to be ethical professionals in a changing landscape of technologies. In: *Hello! Where are you in the landscape of educationaltechnology? Proceedings ascilite Melbourne 2008*.

Frøkjær, E., & Hornbæk, K. (2008). Metaphors of human thinking for usability inspection and design. *ACM Transactions on Computer-Human Interaction (TOCHI), 14*(4), 20.

Giraud, S., Colombi, T., Russo, A., & Thérouanne, P. (2011). Accessibility of rich internet applications for blind people: a study to identify the main problems and solutions. In: *Proceedings of the 9th ACM SIGCHI Italian Chapter International Conference on Computer-Human Interaction: Facing Complexity*, 163-166.

Gonçalves, R., Martins, J., & Branco, F. (2014). A Review on the Portuguese Enterprises Web Accessibility Levels–A website accessibility high level improvement proposal. *Procedia Computer Science, 27*, 176-185.

Gonçalves, R., Martins, J., Pereira, J., Oliveira, M. A.-Y., & Ferreira, J. J. P. (2012). Accessibility levels of Portuguese enterprise websites: equal opportunities for all? *Behaviour & Information Technology, 31*(7), 659-677.

Gonçalves, R., Martins, J., Pereira, J., Oliveira, M. A.-Y., & Ferreira, J. J. P. (2013). Enterprise Web accessibility levels amongst the Forbes 250: Where art thou o virtuous leader? *Journal of business ethics, 113*(2), 363-375.

Good, A. (2008). *An investigation of a method for improving accessibility to web-based information for users with impairments.* PhD thesis, University of Portsmouth.

Grantham, J., Grantham, E., & Powers, D. (2012). Website accessibility: an Australian view. In: *Proceedings of the Thirteenth Australasian User Interface Conference-Volume 126*, 21-28.

Hackett, S., & Parmanto, B. (2009). Homepage not enough when evaluating web site accessibility. *Internet Research, 19*(1), 78-87.

Hackett, S., Parmanto, B., & Zeng, X. (2004). Accessibility of Internet websites through time. In: *ACM SIGACCESS Accessibility and Computing*, 32-39.

Harper, S., & Chen, A. Q. (2012). Web accessibility guidelines. *World Wide Web, 15*(1), 61-88.

Harper, S., & Yesilada, Y. (2008). *Web accessibility: a foundation for research*: Springer Science & Business Media.

Hart, S. G. (2006). NASA-task load index (NASA-TLX); 20 years later. In: *Proceedings of the human factors and ergonomics society annual meeting*, 904-908.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in psychology, 52*, 139-183.

Hassanzadeh, M., & Navidi, F. (2010). Web site accessibility evaluation methods in action: A comparative approach for ministerial web sites in Iran. *The Electronic Library, 28*(6), 789-803.

Hassenzahl, M. (2004). The interplay of beauty, goodness, and usability in interactive products. *Human-computer interaction, 19*(4), 319-349.

Hassenzahl, M., & Tractinsky, N. (2006). User experience-a research agenda. *Behaviour & Information Technology, 25*(2), 91-97.

Henry, S. L. (2007). *Just ask: integrating accessibility throughout design*: Lulu.com.

Herring, S. C., Scheidt, L. A., Wright, E., & Bonus, S. (2005). Weblogs as a bridging genre. *Genres of Digital Documents, 18*(2), 142-171.

Hertzum, M., Hansen, K. D., & Andersen, H. H. (2009). Scrutinising usability evaluation: does thinking aloud affect behaviour and mental workload? *Behaviour & Information Technology, 28*(2), 165-181.

Hertzum, M., & Jacobsen, N. E. (2001). The evaluator effect: A chilling fact about usability evaluation methods. *International Journal of Human-Computer Interaction, 13*(4), 421-443.

Hoffman, D. L., Novak, T. P., & Venkatesh, A. (2004). Has the Internet become indispensable? *Communications of the ACM, 47*(7), 37-42.

Holsapple, C. W., Pakath, R., & Sasidharan, S. (2005). A website interface design framework for the cognitively impaired: A study in the context of alzheimer's disease. *Journal of Electronic Commerce Research, 6*(4), 291.

Horton, S., & Quesenbery, W. (2014). *Web for Everyone: designing accessible user experiences*: Rosenfeld Media.

Horton, S., & Sloan, D. (2014). Accessibility in practice: a process-driven approach to accessibility. In: *Inclusive Designing*, 105-115.

Howell, D. C. (2012). *Statistical methods for psychology*: Cengage Learning.

Huffaker, R. (2014). The United States' legislative impact on eaccessibility: what the European Union can learn. In: *Proceedings of the 11th Web for All Conference*, 14.

Iaccarino, G., Malandrino, D., & Scarano, V. (2006). Personalizable edge services for web accessibility. In: *Proceedings of the 2006 international cross-disciplinary workshop on Web accessibility*

(W4A): Building the mobile web: rediscovering accessibility?, 23-32.

Isaila, N., & Nicolau, I. (2010). Promoting computer assisted learning for persons with disabilities. *Procedia-Social and Behavioral Sciences, 2*(2), 4497-4501.

ISO 9241-171. (2008). Ergonomics of human-system interaction-Part 171: Guidance on software accessibility. *Geneva: International Standards Organization.*

ISO 9241-210. (2010). Ergonomics of human system interaction-Part 210: Human-centred design for interactive systems. *International Standardization Organization (ISO). Switzerland*.

Iwata, H., Kobayashi, N., Tachibana, K., Shirogane, J., & Fukazawa, Y. (2013). Web accessibility support for visually impaired users using link content analysis. *SpringerPlus, 2*(1), 116.

Jaeger, P. T. (2006). Assessing Section 508 compliance on federal e-government Web sites: A multi-method, user-centered evaluation of accessibility for persons with disabilities. *Government Information Quarterly, 23*(2), 169-190.

Jaeger, P. T. (2008). User-centered policy evaluations of section 508 of the rehabilitation act: Evaluating e-government web sites for accessibility for persons with disabilities. *Journal of Disability Policy Studies*, 24-33.

Jati, H., & Dominic, D. D. (2008). Website accessibility performance evaluation in Malaysia. In: *Information Technology, 2008. ITSim 2008. International Symposium on*, 1-3.

Jitaru, E., & Alexandru, A. (2008). Content Accessibility of Web documents. Principles and Recommendations. *Informatica Economica Journal, 12*(2), 117-124.

Joaquín Mira, J., Llinás, G., Tomás, O., & Pérez-Jover, V. (2006). Quality of websites in Spanish public hospitals. *Medical Informatics and the Internet in Medicine, 31*(1), 23-44.

Johnson, F., Rowley, J., Craven, J., Johnson, F., & Butters, G. (2010). The usability and functionality of an online catalogue. In: *Aslib Proceedings*, 70-84.

Jørgensen, A. H. (1990). Thinking-aloud in user interface design: a method promoting cognitive ergonomics. *Ergonomics, 33*(4), 501-507.

Kane, D., & Hegarty, N. (2007). New Web site, new opportunities: Enforcing standards compliance within a content management system. *Library hi tech, 25*(2), 276-287.

Kane, S. K. (2007). Everyday Inclusive Web Design: An Activity Perspective. *Information Research, 12*(1).

Karkin, N., & Janssen, M. (2014). Evaluating websites from a public value perspective: A review of Turkish local government websites. *International Journal of Information Management, 34*(3), 351-363.

Kerkmann, F., & Lewandowski, D. (2012). Accessibility of web search engines: Towards a deeper understanding of barriers for people with disabilities. *Library Review, 61*(8/9), 608-621.

Kern, W. (2008). Web 2.0-end of accessibility? analysis of most common problems with Web 2.0 based applications regarding

Web accessibility. *International Journal of Public Information Systems, 4*(2).

Kituyi, G. M., & Anjoga, H. (2013). Improvement of e-Government Service Usability in Developing Countries: Empirical Experiences of Uganda. *Journal of Emerging Trends in Computing and Information Sciences, 4*(3), 263-269.

Krippendorff, K. (2004). *Content analysis: An introduction to its methodology*: Sage.

Kurt, S. (2012). Issues to consider in designing WebQuests: A literature review. *Computers in the Schools, 29*(3), 300-314.

Kurt, S. (2013). Accessibility Issues of Educational Web Sites. *Assistive Technologies: Concepts, Methodologies, Tools, and Applications: Concepts, Methodologies, Tools, and Applications*, 54-62.

Kuzma, J. M. (2010). Accessibility design issues with UK e-government sites. *Government Information Quarterly, 27*(2), 141-146.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *biometrics, 33*, 159-174.

Latif, M. H. A., & Masrek, M. N. (2010). Accessibility evaluation on malaysian e-Government websites. *Journal of e-Government Studies and Best Practice*.

Lavery, D., Cockton, G., & Atkinson, M. P. (1997). Comparison of evaluation methods using structured usability problem reports. *Behaviour & Information Technology, 16*(4-5), 246-266.

Lazar, J., Beavan, P., Brown, J., Coffey, D., Nolf, B., Poole, R., Turk, R., Waith, V., Wall, T., & Weber, K. (2010). Investigating the accessibility of state government web sites in Maryland. *Designing Inclusive Interactions*, 69-78.

Lazar, J., Beere, P., Greenidge, K.-D., & Nagappa, Y. (2003). Web accessibility in the Mid-Atlantic United States: a study of 50 homepages. *Universal Access in the Information Society, 2*(4), 331-341.

Lazar, J., Feng, J., & Allen, A. (2006). Determining the impact of computer frustration on the mood of blind users browsing the web. In: *Proceedings of the 8th international ACM SIGACCESS conference on Computers and accessibility*, 149-156.

Lazar, J., Olalere, A., & Wentz, B. (2012). Investigating the accessibility and usability of job application web sites for blind users. *Journal of Usability Studies, 7*(2), 68-87.

Lazar, J., Schroeder-Thomas, C., Jones, A., Greenidge, K., Beere, P., & Clements, J. (2003). Detour Ahead: Current Roadblocks to Web Accessibility. In: *Proceedings of the 2003 Conference on Universal Access in Human-Computer Interaction*, 990-994.

Lazar, J., Wentz, B., Almalhem, A., Catinella, A., Antonescu, C., Aynbinder, Y., Bands, M., Bastress, E., Chan, B., & Chelden, B. (2013). A longitudinal study of state government homepage accessibility in Maryland and the role of web page templates for improving accessibility. *Government Information Quarterly, 30*(3), 289-299.

Lazar, J., Wentz, B., Bogdan, M., Clowney, E., Davis, M., Guiffo, J., Gunnarsson, D., Hanks, D., Harris, J., Holt, B., Kitchin, M., Motaynem, M., Nzokou, R., Sedaghat, L., & Stern, K. (2011). Potential pricing discrimination due to inaccessible web sites. In:

Human-Computer Interaction – INTERACT 2011 Lecture Notes in Computer Science, 108-114.

Lee, S., Kim, B. G., & Kim, J. G. (2007). Accessibility evaluation of Korean e-government. In: *International Conference on Universal Access in Human-Computer Interaction*, 73-78.

Leporini, B., Buzzi, M. C., & Buzzi, M. (2010). Providing web accessibility for the visually impaired. *ERCIM News, 80*, 32-33.

Leporini, B., & Paternò, F. (2008). Applying web usability criteria for vision-impaired users: does it really improve task performance? *International Journal of Human-Computer Interaction, 24*(1), 17-47.

Letourneau, C. (1998). Accessible Web design - a definition. *Starling Access Services.* Retrieved from http://www.starlingweb.com/webac.htm

Lewis, C. (1982). *Using the" thinking-aloud" method in cognitive interface design*: IBM TJ Watson Research Center.

Li, S.-H., Yen, D. C., Lu, W.-H., & Lin, T.-L. (2012). Migrating from WCAG 1.0 to WCAG 2.0–A comparative study based on Web Content Accessibility Guidelines in Taiwan. *Computers in Human Behavior, 28*(1), 87-96.

Lopes, R., & Carrico, L. (2010). Macroscopic characterisations of Web accessibility. *New Review of Hypermedia and Multimedia, 16*(3), 221-243.

Lorca, P., Andrées, J. D., & Martínez, A. B. (2012). Size and culture as determinants of the web policy of listed firms: The case of web accessibility in Western European countries. *Journal of the*

American society for Information science and technology, 63(2), 392-405.

Luján-Mora, S., & Masri, F. (2012). Evaluation of web accessibility: A combined method. *Information Systems Research and Exploring Social Artifacts: Approaches and Methodologies, Information Science Reference*, 314-331.

Luján-Mora, S., Navarrete, R., & Peñafiel, M. (2014). Egovernment and web accessibility in South America. In: *First International Conference on eDemocracy & eGovernment (ICEDEG)* 77-82.

Maatta Smith, S. L. (2014). Web accessibility assessment of urban public library websites. *Public Library Quarterly, 33*(3), 187-204.

Mancini, C., Zedda, M., & Barbaro, A. (2005). Health information in Italian public health websites: moving from inaccessibility to accessibility. *health information & libraries journal, 22*(4), 276-285.

Mankoff, J., Fait, H., & Tran, T. (2005). Is your web page accessible?: a comparative study of methods for assessing web page accessibility for the blind. In: *Proceedings of the SIGCHI conference on Human factors in computing systems*, 41-50.

Márquez, S., Moreno, F., Coret, J., Jiménez, E., Alcantud, F., & Guarinos, I. (2012). Web accessibility for visual disabled: an expert evaluation of the Inclusite® solution. In: *Interactive Collaborative Learning (ICL), 2012 15th International Conference on*, 1-5.

Martín, A., Cechich, A., & Rossi, G. (2008). Comparing approaches to web accessibility assessment. *Handbook of research on web information systems quality*, 181-205.

Martin, A., Mazalu, R., & Cechich, A. (2010). Supporting an aspect-oriented approach to web accessibility design. In: *Fifth International Conference on Software Engineering Advances (ICSEA),* , 20-25.

Martínez, A. B., De Andrés, J., & García, J. (2014). Determinants of the Web accessibility of European banks. *Information Processing & Management, 50*(1), 69-86.

Martins, J., Cruz, J., & Gonçalves, R. (2009). Portuguese Web Accessibility-Portuguese Enterprises Websites Accessibility Evaluation. In: *ICEIS (5)*, 142-147.

Maswera, T., Dawson, R., & Edwards, J. (2005). Analysis of usability and accessibility errors of e-commerce websites of tourist organisations in four African countries. *Information and Communication Technologies in Tourism 2005*, 531-542.

McDonald, S., Edwards, H. M., & Zhao, T. (2012). Exploring think-alouds in usability testing: An international survey. *IEEE transactions on professional communication, 55*(1), 2-19.

McDonald, S., Zhao, T., & Edwards, H. M. (2015). Look Who's Talking: Evaluating the Utility of Interventions During an Interactive Think-Aloud. *Interacting with computers, 28*(3), 387–403.

Mereuţă, A., Aupetit, S., & Slimane, M. (2012). Improving web accessibility for dichromat users through contrast preservation.

In: *International Conference on Computers for Handicapped Persons*, 363-370.

Michopoulou, E., & Buhalis, D. (2013). Information provision for challenging markets: The case of the accessibility requiring market in the context of tourism. *Information & Management, 50*(5), 229-239.

Mills, J. E., Han, J.-H., & Clay, J. M. (2008). Accessibility of hospitality and tourism websites a challenge for visually impaired persons. *Cornell Hospitality Quarterly, 49*(1), 28-41.

Mitsamarn, N., Gestubtim, W., & Junnatas, S. (2007). Web accessibility: a government's effort to promote e-accessibility in Thailand. In: *Proceedings of the 1st international convention on Rehabilitation engineering & assistive technology: in conjunction with 1st Tan Tock Seng Hospital Neurorehabilitation Meeting*, 23-27.

Modesto, D. M., & Ferreira, S. B. L. (2014). Guidelines for search features development–a comparison between general users and users with low reading skills. *Procedia Computer Science, 27*, 334-342.

Moreno, L., Martínez, P., & Ruiz-Mezcua, B. (2009). A bridge to web accessibility from the usability heuristics. In: *Symposium of the Austrian HCI and Usability Engineering Group*, 290-300.

Mori, G., Buzzi, M. C., Buzzi, M., Leporini, B., & Penichet, V. M. (2011). Collaborative editing for all: the google docs example. In: *International Conference on Universal Access in Human-Computer Interaction*, 165-174.

Morville, P. (2005). Experience design unplugged. In: *ACM SIGGRAPH 2005 Web program*, 10.

Nielsen, J. (1994). *Usability engineering*: Elsevier.

Norman, D. A. (1988). *The psychology of everyday things*: Basic books.

Norman, D. A. (2013). *The design of everyday things: Revised and expanded edition*: Basic books.

Noyes, J. M., & Bruneau, D. P. (2007). A self-analysis of the NASA-TLX workload measure. *Ergonomics, 50*(4), 514-519.

Olalere, A., & Lazar, J. (2011). Accessibility of US federal government home pages: Section 508 compliance and site accessibility statements. *Government Information Quarterly, 28*(3), 303-309.

Olmsted-Hawala, E. L., Murphy, E. D., Hawala, S., & Ashenfelter, K. T. (2010). Think-aloud protocols: a comparison of three think-aloud protocols for use in testing data-dissemination web sites for usability. In: *Proceedings of the sigchi conference on human factors in computing systems*, 2381-2390.

Oppenheim, C., & Selby, K. (1999). Access to information on the World Wide Web for blind and visually impaired people. In: *Aslib Proceedings*, 335-345.

Paciello, M. (1996a). Making the Web accessible for the deaf, hearing and mobility impaired. *Florida Libraries, 39*(5), 83-91.

Paciello, M. (1996b). Making the world wide web accessible for the blind and visually impaired. *Florida Libraries, 39*(5).

Paciello, M. (1996c). The web and people with disabilities: Cutting edge developments. *Florida Libraries, 39*(5).

Paris, M. (2006). Website accessibility: a survey of local e-government websites and legislation in Northern Ireland. *Universal Access in the Information Society, 4*(4), 292-299.

Park, E.-J., & Lim, H. (2014). A Study on the Evaluation of the Web Accessibility of a Domestic General Shopping Site with a High Number of Visits: Focusing on the Attributes of Alternative Texts with< img> Tags. *Current Research Trend of Multimedia, Signal Processing and Software*, 17-20.

Park, E.-J., Lim, Y.-W., & Lim, H.-K. (2014). A study of web accessibility of websites built in html5-focusing on the top 100 most visited websites. *International Journal of Multimedia and Ubiquitous Engineering, 9*(4), 247-256.

Parkinson, C., & Olphert, C. (2010). Website Accessibility and the Role of Accessibility Statements. *Integrating Usability Engineering for Designing the Web Experience: Methodologies and Principles*, 166-190.

Patch, K., Spellman, J., & Wahlbin, K. (2015). Mobile Accessibility: How WCAG 2.0 and Other W3C/WAI Guidelines Apply to Mobile. Retrieved from https://www.w3.org/TR/mobile-accessibility-mapping/

Pereira, L. S., Ferreira, S. B. L., Braga, H., de Castro Salgado, L. C., & Nunes, R. R. (2014). Using Cultural Viewpoint Metaphors to Provide Web Accessibility for the Visually Impaired Users. *Procedia Computer Science, 27*, 186-196.

Peters, C., & Bradbard, D. A. (2007). Web accessibility: an introduction and implications for a corporate social responsibility marketing strategy. *Journal of Internet Commerce, 6*(4), 27-54.

Petrie, H., & Kheir, O. (2007). The relationship between accessibility and usability of websites. In: *Proceedings of the SIGCHI conference on Human factors in computing systems*, 397-406.

Petrie, H., & Power, C. (2012). What do users really care about?: a comparison of usability problems found by users and experts on highly interactive websites. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2107-2116.

Petrie, H., & Precious, J. (2010). Measuring User Experience of websites: Think aloud protocols and an emotion word prompt list. In: *CHI'10 Extended Abstracts on Human Factors in Computing Systems*, 3673-3678.

Power, C., Freire, A., Petrie, H., & Swallow, D. (2012). Guidelines are only half of the story: accessibility problems encountered by blind users on the web. In: *Proceedings of the SIGCHI conference on human factors in computing systems*, 433-442.

Power, C., Petrie, H., Freire, A. P., & Swallow, D. (2011). Remote evaluation of WCAG 2.0 techniques by web users with visual disabilities. In: *International Conference on Universal Access in Human-Computer Interaction*, 285-294.

Power, C., Petrie, H., Sakharov, V., & Swallow, D. (2010). Virtual learning environments: another barrier to blended and e-learning. In: *International Conference on Computers for Handicapped Persons*, 519-526.

Power, C., Petrie, H., Swallow, D., Murphy, E., Gallagher, B., & Velasco, C. A. (2013). Navigating, Discovering and Exploring the Web: Strategies Used by People with Print Disabilities on Interactive

Websites. In: *Human-Computer Interaction – INTERACT 2013*, 667-684.

Pribeanu, C., Gheorghe-Moisii, M., & Fogarassy-Neszly, P. (2015). Accessibility of Romanian Municipal Websites–Conformance with WCAG2. In: *Proceedings of RoCHI*, 25-28.

Ramayah, B., Jaafar, A., & Yatim, N. F. M. (2013). Visually Impaired User's Navigation Experiences in Facebook. In: *International Visual Informatics Conference*, 788-796.

Reid, L. G., & Snow-Weaver, A. (2008). WCAG 2.0: a web accessibility standard for the evolving web. In: *Proceedings of the 2008 international cross-disciplinary conference on Web accessibility (W4A)*, 109-115.

Rogers, Y., Sharp, H., & Preece, J. (2007). *Interaction design: beyond human-computer interaction*: John Wiley & Sons.

Rømen, D., & Svanæs, D. (2008). Evaluating web site accessibility: validating the WAI guidelines through usability testing with disabled users. In: *Proceedings of the 5th Nordic conference on Human-computer interaction: building bridges*, 535-538.

Rømen, D., & Svanæs, D. (2012). Validating WCAG versions 1.0 and 2.0 through usability testing with disabled users. *Universal Access in the Information Society, 11*(4), 375-385.

Rubin, J. (1994). Handbook of Usability Testing: How to Plan, Design, and Conduct Effective Tests. *New York: John Wiley & Sons.*

Russo, J. E., Johnson, E. J., & Stephens, D. L. (1989). The validity of verbal protocols. *Memory & cognition, 17*(6), 759-769.

Rutter, R., Lauke, P. H., Waddell, C., Thatcher, J., Henry, S. L., Lawson, B., Kirkpatrick, A., Heilmann, C., Burks, M. R., & Regan, B. (2007). *Web accessibility: Web standards and regulatory compliance*: Apress.

Sambhanthan, A., & Good, A. (2013). Implications for Improving Accessibility to E-Commerce Websites in Developing Countries-A Study of Hotel Websites. *International Journal of Knowledge Based Organizations, 1*(2).

Sánchez-Gordón, M.-L., & Moreno, L. (2014). Toward an integration of Web accessibility into testing processes. *Procedia Computer Science, 27*, 281-291.

Santana, V. F., & Baranauskas, M. C. C. (2014). Log-Based Personalization Tool as an Assistive Technology. In: *International Conference on Human-Computer Interaction*, 433-444.

Section 508. (1996). U.S. Government. Retrieved from https://www.section508.gov/

Section 508 of the Rehabilitation Act. (29 U.S.C. 794d).

Shi, Y. (2007). The accessibility of Chinese local government web sites: An exploratory study. *Government Information Quarterly, 24*(2), 377-403.

Shneiderman, B. (2000). Universal usability. *Communications of the ACM, 43*(5), 84-91.

Shneiderman, B. (2002). Promoting universal usability with multi-layer interface design. In: *ACM SIGCAPH Computers and the Physically Handicapped*, 1-8.

Sierkowski, B. (2002). Achieving web accessibility. In: *Proceedings of the 30th annual ACM SIGUCCS conference on User services*, 288-291.

Slatin, J. M., & Rush, S. (2002). *Maximum accessibility: making your web site more usable for everyone*: Addison-Wesley Longman Publishing Co., Inc.

Sloan, D., & Kelly, B. (2011). Web accessibility metrics for a post digital world. In: *RDWG Symposium on Website Accessibility Metrics*.

Stenitzer, M., Putzhuber, M., Nemecek, S., & Büchler, F. (2008). Accessible Online Shops for the Older Generation and People with Disabilities. *Computers Helping People with Special Needs*, 462-465.

Strain, P., Shaikh, A. D., & Boardman, R. (2007). Thinking but not seeing: think-aloud for non-sighted users. In: *CHI'07 Extended Abstracts on Human Factors in Computing Systems*, 1851-1856.

Suh, C.-K., & CHO, E. (2007). Web content accessibility of disabled community sites in Korea. In: *7th WSEAS International Conference on Applied Computer Science*, 327-333.

Swierenga, S. J., Sung, J., Pierce, G. L., & Propst, D. B. (2011). Website design and usability assessment implications from a usability study with visually impaired users. In: *International Conference on Universal Access in Human-Computer Interaction*, 382-389.

Talib, M. N., Shuqin, C., Abrar, M., & Shafiq, M. S. (2009). E-Business Access for Blinds: A Semantic Approach. In: *International*

Conference on E-Business and Information System Security, 2009. EBISS'09, 1-4.

Thatcher, J., Bohman, P., Burks, M., Henry, S. L., Regan, B., Swierenga, S., Urban, M. D., & Waddell, C. (2002). *Constructing accessible web sites*: Glasshaus Apress.

Theofanos, M. F., & Redish, J. G. (2003). Bridging the gap: between accessibility and usability. *interactions, 10*(6), 36-51.

Thompson, T., Burgstahler, S., & Comden, D. (2003). Research on web accessibility in higher education. *Journal of Information Technology and Disabilities, 9*(2).

Trewin, S., Cragun, B., Swart, C., Brezin, J., & Richards, J. (2010). Accessibility challenges and tool features: an IBM Web developer perspective. In: *Proceedings of the 2010 international cross disciplinary conference on web accessibility (W4A)*, 32.

Van den Haak, M., De Jong, M., & Jan Schellens, P. (2003). Retrospective vs. concurrent think-aloud protocols: testing the usability of an online library catalogue. *Behaviour & Information Technology, 22*(5), 339-351.

Van den Haak, M., De Jong, M., & Jan Schellens, P. (2004). Employing think-aloud protocols and constructive interaction to test the usability of online library catalogues: a methodological comparison. *Interacting with computers, 16*(6), 1153-1170.

Van den Haak, M., De Jong, M., & Jan Schellens, P. (2007). Evaluation of an informational web site: three variants of the think-aloud method compared. *Technical Communication, 54*(1), 58-71.

Van den Haak, M., De Jong, M., & Jan Schellens, P. (2009). Evaluating municipal websites: A methodological comparison of three think-aloud variants. *Government Information Quarterly, 26*(1), 193-202.

Vigo, M., & Harper, S. (2013). Coping tactics employed by visually disabled users on the web. *International Journal of Human-Computer Studies, 71*(11), 1013-1025.

Waddell, C. D. (1998). Applying the ADA to the Internet: A web accessibility standard. Retrieved from http://www.icdri.org/CynthiaW/applying_the_ada_to_the_internet.htm

WAI. (2005). Introduction to web accessibility. Retrieved from http://www.w3.org/WAI/intro/accessibility.php

Wang, M., Liu, B., & Hua, X.-S. (2010). Accessible image search for colorblindness. *ACM Transactions on Intelligent Systems and Technology (TIST), 1*(1), 1-26.

Wanniarachchi, H., & Jayathilake, D. (2012). A framework for building web sites that are friendly to visually impaired. In: *International Conference on Advances in ICT for Emerging Regions (ICTer)*, 103-110.

Watanabe, M., Asai, D., & Asano, Y. (2007). Improving accessibility through the visual structure of web contents. In: *International Conference on Universal Access in Human-Computer Interaction*, 185-192.

Watanabe, T. (2009). Experimental evaluation of usability and accessibility of heading elements. *Disability and Rehabilitation: Assistive Technology, 4*(4), 236-247.

WebAIM. (2014). Screen Reader User Survey #5 Results. Retrieved from http://webaim.org/projects/screenreadersurvey5/

WebAIM. (2017). Alternative Text. Retrieved from http://webaim.org/techniques/alttext/

Wentz, B., Lazar, J., Stein, M., Gbenro, O., Holandez, E., & Ramsey, A. (2014). Danger, danger! Evaluating the accessibility of Web-based emergency alert sign-ups in the northeastern United States. *Government Information Quarterly, 31*(3), 488-497.

Wijayaratne, A., & Singh, D. (2010). Is there space in cyberspace for distance learners with special needs in Asia? A review of the level of Web accessibility of institutional and library homepages of AAOU members. *The International Information & Library Review, 42*(1), 40-49.

Williams, R., Rattray, R., & Grimes, A. (2006). Meeting the on-line needs of disabled tourists: an assessment of UK-based hotel websites. *International Journal of Tourism Research, 8*(1), 59-73.

World Health Organization. (2016). Disability and health. Retrieved from http://www.who.int/mediacentre/factsheets/fs352/en/

World Health Organization. (2017). Visual impairment and blindness. Retrieved from http://www.who.int/mediacentre/factsheets/fs282/en/

Wright, P. C., & Monk, A. F. (1991). The use of think-aloud evaluation methods in design. *ACM SIGCHI Bulletin, 23*(1), 55-57.

Wroblewski, L. (2011). *Mobile First Foreword By Jeffrey Zeldman*.

Yao, D., Qiu, Y., Huang, H., Du, Z., & Ma, J. (2011). A survey of technology accessibility problems faced by older users in China. *Universal Access in the Information Society, 10*(4), 373-390.

Yates, R. (2005). Web site accessibility and usability: towards more functional sites for all. *Campus-wide Information systems, 22*(4), 180-188.

Yesilada, Y., Brajnik, G., & Harper, S. (2011). Barriers common to mobile and disabled web users. *Interacting with computers, 23*(5), 525-542.

Yoon, K., Newberry, T., Hulscher, L., & Dols, R. (2013). Call for library websites with a separate information architecture for visually impaired users. In: *Proceedings of the American Society for Information Science and Technology*, 1-3.

Zeng, X. (2004). *Evaluation and enhancement of web content accessibility for persons with disabilities.* PhD thesis, University of Pittsburgh.

Zerfass, A., & Hartmann, B. (2005). The usability factor: Improving the quality of e-content. *E-Content*, 165-182.